nature

www.nature.com/nature

Vol 446 | Issue no. 7133 | 15 March 2007

In praise of the 'brain drain'

Countries and professions that export skilled staff do not always lose out.

Governments and scientists are often heard expressing angst about the malign impact on their countries of the 'brain drain', the flow of skilled individuals to foreign climes. The brain drain worries people everywhere, with the possible exception of California. It is a policy fixation in European science, a concern for at least three-quarters of American states, and, most of all, a major strategic headache for developing countries.

In South Africa, for example, the government has demonized institutions that train doctors and nurses who leave for employment elsewhere. It has taken steps to penalize those state-trained health professionals who choose to leave.

But South Africa is wrong. Its perception of the brain drain — as a simple transaction in which the recipient gains and the donor loses — is, at best, incomplete. Conventional wisdom holds that the movement of trained healthcare personnel from Africa is creating a crisis in public health. But is that really what is happening?

Michael Clemens, an economist at the non-partisan Center for Global Development in Washington DC, doesn't think so. He reported his findings at the meeting of the American Association for the Advancement of Science in San Francisco last month (see www.cgdev.org/content/publications/detail/13123). There is a clear correlation between emigration and the state of the public healthcare system, but not the one you might expect. The higher the proportion of an African nation's nurses and doctors who have moved abroad, the better shape its healthcare is likely to be in.

This is not so strange, when you think about it. Countries and professions with more openness and greater mobility of personnel are more likely to be in touch with global trends — and more likely to attract able trainees in the first place. The worst public healthcare systems, Clemens says, are in French-speaking West Africa, where staff are least likely to emigrate, as France won't let them.

The tendency of perhaps half of today's emigrants to return home later on in their careers is another factor. So is remuneration and the large amounts of cash that migrants send back home. These changes make the old model of immigrant 'donor' societies obsolescent. Communities can benefit, financially and intellectually, from those who have left. It is the degree to which these benefits counteract the unquestionable initial loss that is open to question.

Similar observations could be made regarding emigration flows between wealthy nations. According to the World Bank, Britain has more professional émigrés than any nation on Earth. But it doesn't seem to be hurting. California's research labs may be crawling with Brits, yet UK science has gone from strength to strength. According to surveys of citations against expenditure, Britain has one of the most productive research systems in the world. How can this be?

Well, say the revisionists, science departments at British universities may actually benefit from the ambition to depart, and, to a lesser degree, from their connections with those who have done so. Perish the thought, but some of these mobile researchers may even do the best work of their lives at Salford, say, only to take their foot ever-so-slightly off the gas when they 'arrive' at Stanford.

Woody Allen once observed that the sole cultural advantage of California (over his native New York, presumably) was its law permitting you to turn right at a red light. To be fair, science and industry in the Golden State have clearly benefited to a massive degree from immigrant talent gleaned from every corner of the planet. But the notion that other places have necessarily suffered a corresponding loss — or that emigration is a zero-sum game — is misplaced. ■

The legacy of Linnaeus

Taxonomy in an age of transformation.

Every plant and animal has a mitochondrial cytochrome oxidase I gene, and its sequence helps researchers assign that plant or animal to a given species, with some degree of certainty. The precise degree of the certainty obtained using this 'barcode' sequence is a matter of some debate, but such sequences are clearly useful to both taxonomists and those who use applied taxonomy. And the industrial-scale sequencing that allowed Craig Venter's ocean-metagenomics consortium to deposit billions of letters of sequence from hundreds of thousands of microbe genes into the GenBank database this week opens up even more possibilities.

The ability to peer into living things and inspect the evolutionary scorecard encoded in their genes has transformed the whole of biology, but few fields have had their core assumptions challenged as deeply as taxonomy. From the time of Carl Linnaeus, born 300 years ago this May, taxonomy has relied on the observation and comparison of physical forms. Now it is supplemented by access to what would once have been seen not as form, but as essence.

Linnaeus himself sought a universal classification of all creation, animal, vegetable and mineral. His categorizations were not uniformly valuable, but his systematic spirit, his stress on the concept of species, and the formal but adaptable conventions of nomenclature he introduced have endured. *Nature* is glad to celebrate his legacy in this special issue.

DNA sequencing is a gift that Linnaeus would surely have made great use of, but it brings its own problems. It is not always easily reconciled with the careful description, annotation and curation that have been the duty and delight of the taxonomists who carried the linnaean programme forward. The availability of DNA sequences

invites both pure scientists (see page 247) and conservationists in the field (see page 250) to change their ways of working.

The classical world in which Linnaeus worked may seem, at first glance, to contrast with our present age of change. Linnaeus believed in fixed species of knowable number created by God and observable by men, in a world more like the lawns and flowerbeds of a formal garden than Darwin's dynamic "tangled bank".

Yet Linnaeus's classification was itself a response to a changing world — a world in which Europe's growing hegemony was bringing new species into the realm at headlong speed. Linnaeus's own experience of it was bounded by France to the south and Lapland to the north, but the 'apostles' who carried forth his words and sent back samples and descriptions sailed out from Sweden to Arabia, the Americas north and south, China, Japan and the Pacific. Two sailed with Captain Cook, sending back samples from the parts of the Pacific now being sieved for genes.

The various inventories that Linnaeus produced grew ever longer. But while today's world continues to expand in many ways, in some, including those most important to taxonomy, it is shrinking. The creationist Linnaeus was able to assert that "we can count as many species now as were created at the beginning", but today's taxonomists suspect with near certainty that species are being irretrievably lost to science at an ever-quickening rate.

This whittling away of the bark and marrow of life is not necessarily, in itself, catastrophic. Not every lost species represents a depletion of ecosystem services or other human amenity, and although that is not the only yardstick by which to measure such things, it is in some ways the most important.

Linnaeus would no doubt find much to admire in today's scientific world: its scholarship; its ability to assemble data from around the world in minutes; its tools for examining essences. He would hopefully come round to evolutionary theory — and see the error of the racial categorizations he applied to humans.

The fact that so much of life can be seen in a few buckets of sea water might reconcile him to the fact that swallows do not, as he insisted, wait out the winter in lake-bottom mud. But the realization that the second edition of his dreamed-of universal catalogue would be slimmer than its first would surely strike him as a melancholy one. ■

Open for business

California's stem-cell institute is already transparent enough.

Twenty-eight months after California's electorate voted to spend $3 billion on an institute dedicated to stem-cell research, the California Institute for Regenerative Medicine (CIRM) is finally up and running.

Last month, it gave 72 labs SEED (Scientific Excellence through Exploration and Development) grants for innovative approaches to stem-cell work. Tomorrow they will be joined by up to 25 more, under the institute's first tranche of mainstream investigator grants.

Details of how these grants will be executed — often in the same labs as other research funded by the federal government, and therefore subject to the Bush administration's tight constraints on stem-cell research — have yet to be fully worked out (see page 238).

However, it is already clear that the CIRM has established a robust grant-review mechanism that should satisfy the critics who said the institute would be secretive, biased towards particular organizations that had lobbied for its creation, or overly deferential to scientists. Indeed, the scrutiny to which the CIRM was initially subjected has resulted in a system for grant evaluation that is in some respects more transparent than that of any other research agency.

For example, the CIRM publishes a large amount of information about every grant application on its website (www.cirm.ca.gov), including anonymous reviewers' comments, review scores and the reviewers' recommendations to grant or withhold funding. Successful grant applicants are named on the site, but those who fail remain thankfully anonymous.

The reviewers' comments are frank and potentially embarrassing. One review summary opined that a successful applicant's proposal "lacks focus" and "is not well put together", and some reviewers said they were sceptical about the rigour of another successful applicant's prior publication record.

This openness derives from the CIRM's unusual beginnings. In its short existence, the agency has engaged in constant and often productive dialogue with watchdog groups such as the Foundation for Taxpayer and Consumer Rights, based in Santa Monica, and members of the California legislature in Sacramento. The CIRM's oversight procedures and structure have come under close scrutiny through lawsuits in the California courts.

But as the first grants are awarded, the CIRM is under pressure to open up its review processes still further. The taxpayer-rights group now wants the agency to identify not only those who win awards, but also those who lose out. That would be akin to the state of California publicly releasing information on all the job applications it receives, complete with adverse comments made during the hiring process.

It is impossible to see how such a move will benefit California's taxpayers. Publicly identifying, and sometimes humiliating, those who fail to win awards serves no useful purpose. On the contrary, it is likely to deter scientists from submitting risky proposals that might draw flak from reviewers, and may well curtail reviewers' honesty.

Watchdog groups have also suggested that more financial information should be disclosed about members of the grant-review working group itself. The group is recruited from outside California, and is already screened for potential professional and financial conflicts of interest. The proposal for yet more disclosure, if implemented, would make the CIRM's external review process more onerous, without adding useful information on potential conflicts.

The CIRM is now functioning well, and its dialogue with the public has played a significant role in its progress. Calls for yet more openness may be well intentioned, but they threaten to override the element of confidentiality that is inherent to fair peer review, and to undercut the agency's mission of supporting cutting-edge research from the best Californian scientists. There comes a point at which yet more sunshine leads to sunburn. ■

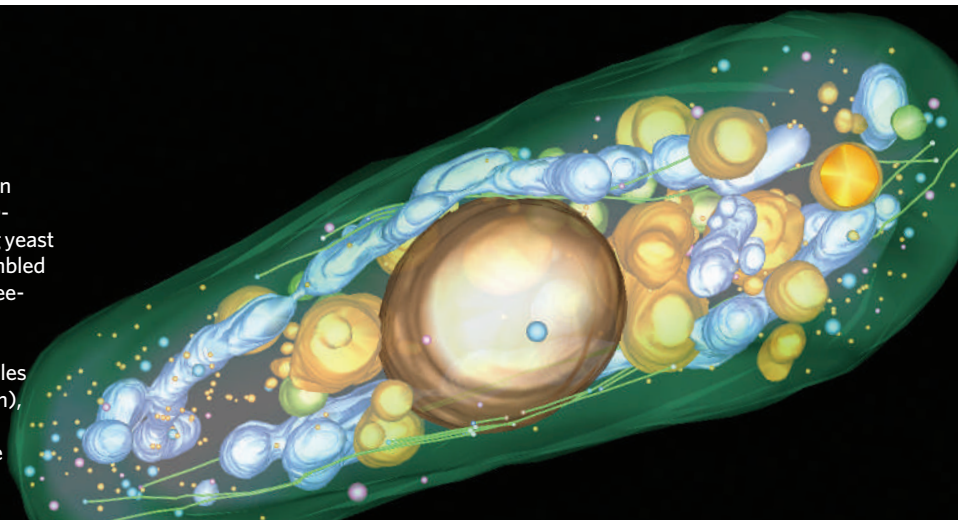
RESEARCH HIGHLIGHTS

An added dimension

Dev. Cell **12,** 349–361 (2007)

This high-resolution, three-dimensional image of a eukaryotic cell was created by researchers led by Claude Antony of the European Molecular Biology Laboratory in Heidelberg, Germany. They scanned 250-nanometre-thick slices of frozen budding yeast (*Schizosaccharomyces pombe*), and assembled the scans into a computer-generated three-dimensional image.

The snapshot here shows the nucleus (coloured brown) and vacuoles and vesicles (gold). Also seen are microtubules (green), which form the cell's 'cytoskeleton', and mitochondria (blue) — most of which are strung along the microtubule bundles.



J. HÖÖG, EMBL

DRUG DELIVERY

Transport in the fast lane

Nature Biotechnol. doi:10.1038/nbt1292 (2007)

Researchers have caught on film a technique that can rapidly transport drugs to cells deep within tissue compartments.

The method targets pockets in cell membranes, called caveolae, which import molecules into a cell and are abundant in blood-vessel linings. Jan Schnitzer of the Sidney Kimmel Cancer Center in San Diego, California, and his colleagues injected live mice with fluorescently labelled antibodies targeted to lung caveolae. Videos revealed that the caveolae actively pumped the antibodies from the blood into lung tissue within seconds. Drugs, nanoparticles or gene-therapy vectors could be attached to such antibodies.

CELL BIOLOGY

Sun sensor

Cell **128,** 853–864 (2007)

The tumour-suppressor protein p53 can boast a new function: regulator of the suntan.

A. SIEGEL

It was already known that p53 governs DNA repair responses to ultraviolet radiation and other stresses. Now, David Fisher of Harvard Medical School in Boston, Massachusetts, and his colleagues have found that p53 also directly stimulates expression of a protein called pro-opiomelanocortin, which is involved in producing the skin pigment melanin.

Mice lacking p53 did not tan in response to either ultraviolet radiation or a drug that normally causes hyperpigmentation. Furthermore, human skin cancers that contained a mutated form of p53 also lacked pigment-producing cells. Taken together, the results indicate that p53 may act as a ultraviolet sensor in the skin that activates tanning.

MATERIALS

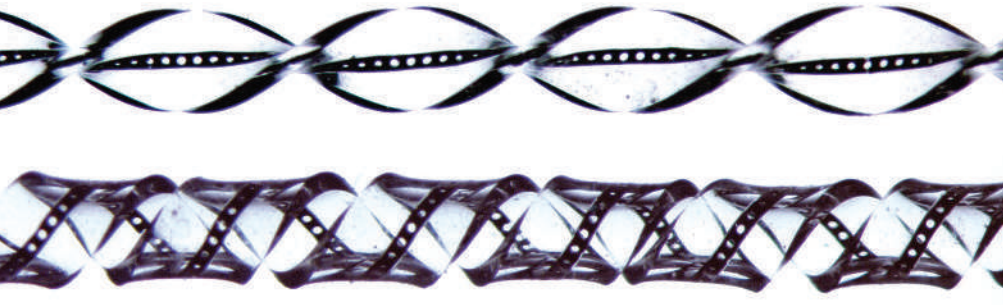
Microsolidics starts here

Adv. Mater. **19,** 727–733 (2007)

Solder would seem a strange thing to inject into microfluidic channels, but George Whitesides and his colleagues at Harvard University in Cambridge, Massachusetts, have created micrometre-wide wires and three-dimensional structures by doing just that. They dub the technique 'microsolidics'.

Micrometre-sized channels in polymer sheets typically serve as conduits for chemicals in microreactors, or for fluids in biological assays. To adapt such channels to carry molten solder, the researchers chemically treated the inside surfaces, oxidizing the polymer then adding a layer of silane.

Wires that formed when the solder solidified could be twisted (pictured below), knotted or woven; re-heating and applying pressure fixed any breaks. The team demonstrated the technique's potential by fabricating a wire that wraps around an open channel; it operates as a heating element.



EARTH SCIENCE

Field reversal

Europhys. Lett. **77,** 59001 (2007)

Clues to how the Earth's magnetic field flips may come from a spinning bathtub-sized cylinder of molten sodium.

Researchers in France have used this experimental set-up to simulate Earth's spinning core of molten iron. Electrical currents in the swirling sodium, hotter than 98 °C, set up a magnetic field analogous to that of Earth.

Michaël Berhanu of the École Normale Supérieure in Paris and his co-workers observed that the sodium's field spontaneously flipped direction every minute or so — mimicking the magnetic field reversals seen when Earth's magnetic poles disappear, then re-emerge upside down. Geological records reveal Earth's field to have reversed sporadically throughout history, but the mechanism remains poorly understood.

STRUCTURAL BIOLOGY

Drug-hunters on target

Science **315,** 1402–1405 (2007)

Antibiotic drug-hunters have long been searching for ways to disrupt a particular enzyme that bacteria use to build their cell walls. Knowing its structure should help.

The enzyme, known as glycosyltransferase, is one of two that can occur as separate domains on the same membrane protein. The other is already the target of antibiotics such as penicillin.

Natalie Strynadka and her colleagues at the University of British Columbia in Vancouver, Canada, have now deciphered the crystal structure of both parts of this membrane protein from *Staphylococcus aureus*. The team also worked out the structure of the protein when it was bound to an antibiotic that's effective in animals and works by blocking glycosyltransferase. That complex might provide clues as to how the antibiotic could be modified to work in humans.

CELL BIOLOGY

Energy from indigestion

Cell **128,** 931–946 (2007)

Cells undergoing apoptosis, or programmed cell death, digest material within themselves to summon up the energy to send out 'clean-up' signals, researchers have shown.

This helps to explain why autophagy — a digestive process activated in normal cells to recycle components such as mitochondria — happens in cells that are destined to die. The finding may also be relevant to understanding diseases associated with cell remnants, such as inflammatory and autoimmune disorders.

Beth Levine and her colleagues at the University of Texas Southwestern Medical Center in Dallas made the discovery by growing mouse embryonic cells under conditions that mimic embryo development, a process that involves massive apoptotic cell death. Apoptotic corpses accumulated when cells lacked either of two autophagy genes, *atg5* or *beclin 1*.

SOLID-STATE PHYSICS

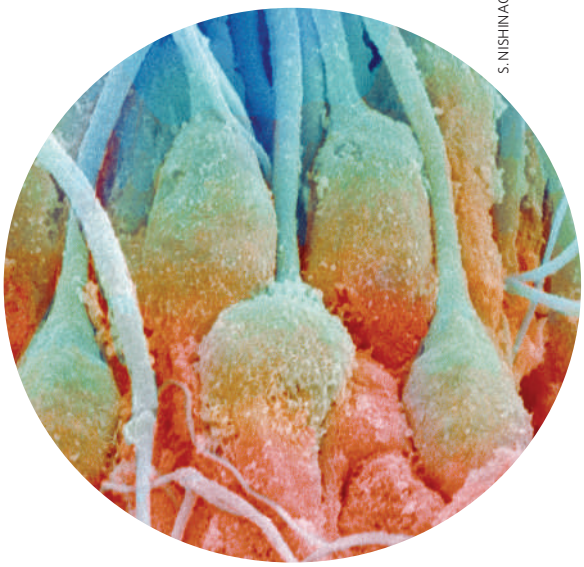
Four thought

Nature Mater. doi:10.1038/nmat1860 (2007)

The first memory device to simultaneously exploit the electric and magnetic properties of a 'multiferroic' material has been demonstrated by researchers in France.

The device was built from a thin layer of lanthanum bismuth manganite, a material described as multiferroic because it has both ferroelectric and ferromagnetic properties. The promise is that memory elements made from such materials could be written electrically and read magnetically — both desirable features for next-generation memory devices.

Manuel Bibes at the Institute of Fundamental Electronics (part of the CNRS, France's basic-research agency) in Orsay and his colleagues have taken a first step by making a device that can adopt four different states. Such memory elements could be used for 'quarternary' logic, rather than the binary logic of conventional memories.



S. NISHINAGA/SPL

REPRODUCTIVE BIOLOGY

RNA reveals male infertility

Hum. Mol. Genet. doi:10.1093/hmg/ddm012 (2007)

Sperm don't just deliver DNA to the egg; they also carry a side-order of RNA. Now, Stephen Krawetz of Wayne State University in Detroit, Michigan, and his colleagues have exploited this RNA to probe the underlying causes of male infertility.

Some infertile men have normal-looking sperm, making it hard to know what the problem is. Others produce sperm with physical abnormalities and other identifiable defects, but even in these cases, the molecular causes are largely unknown.

Krawetz and his team compared the RNA profiles of normal and abnormal sperm. They found marked differences, and showed that many abnormal sperm lacked RNAs involved in processing proteins. The team hopes that the technique will provide diagnoses for men with unexplained infertility.

JOURNAL CLUB

Takuzo Aida
The University of Tokyo, Japan

A chemist sees a commercial future for designer polymers.

Just over a decade ago, a discovery in polymer chemistry triggered explosive progress in macromolecular engineering. A trio of papers published last year will, I think, help to usher the benefits of this development into industry.

The results concern chemists' ability to grow tailored polymers, which have controllable size and architecture. Such polymers are becoming more and more important as major players in the burgeoning field of nanotechnology.

The original breakthrough was the development of a method known as atom-transfer radical polymerization, which made it easy to grow polymers to design.

The method uses a catalyst containing a transition metal, such as copper. The catalyst interacts with the polymer, turning it briefly into a reactive radical that will bind another monomer. Each molecule grows one step at a time, so producing polymers with uniform properties.

A problem with this method has been the large amount of catalyst needed to drive the reaction — leaving residues that are costly to remove. Two recent papers do away with this concern, cutting the concentration of catalyst required by up to 1,000-fold (W. Jakubowski & K. Matyjaszewski *Angew. Chem. Int. Edn* **45,** 4482–4486, 2006; K. Matyjaszewski *et al. Proc. Natl. Acad. Sci. USA* **103,** 15309–15314, 2006).

Another advance, which takes advantage of an unexpectedly active oxidation state of the catalyst's copper, will allow production of polymers with an ultra-high molecular weight and a narrow molecular-weight distribution (V. Percec *et al. J. Am. Chem. Soc.* **128,** 14156–14165, 2006).

Some chemical companies are already setting up industrial plants to make polymers by atom-transfer radical polymerization. These developments mean that more are sure to follow.

is sandwiched at the radical's core between two chemical groups that contain vanadium atoms. This stabilizes the structure, report Christopher Cummins at the Massachusetts Institute of Technology in Cambridge and his team, because the phosphorus shares its lone electron with the two metal atoms. They hope that the same strategy will stabilize radicals centred on other elements, and that varying the metal will tune the radical's reactivity.

MALARIA

Which mozzies win out?

Proc. Natl Acad. Sci. USA **104,** 5580–5583 (2007)

Genetic resistance to the malaria parasite gives mosquitoes feeding on infected blood a fitness advantage, researchers have found. Release of mosquitoes that are resistant to infection with malaria is one control strategy being considered to curb the disease.

Marcelo Jacobs-Lorena of the Johns Hopkins University in Baltimore, Maryland, and his colleagues put 250 transgenic and 250 wild-type mosquitoes of opposite sexes into a cage, where they fed on mice infected with the *Plasmodium berghei* parasite. The transgene, which blocks infection through the mosquitoes' gut, was found in around 70% of the mosquito population after 10 or so breeding cycles. Mosquitoes in the wild only occasionally become infected with the parasite, but this study gives hope that the transgene could persist in the population.

ASTRONOMY

Seeing things

Astrophys. J. **657,** 669–680 (2007)

Infrared light thought to have been emitted by the Universe's first stars isn't seen in a new survey of the skies.

Researchers had previously found a 'near-infrared background excess' in some

NASA/ESA/R. THOMPSON

satellite images that they couldn't account for with known sources. They argued that it was light from early galaxies, stretched by the expansion of the Universe to appear at infrared wavelengths.

Rodger Thompson of the University of Arizona, Tucson, and his colleagues analysed sharper and more sensitive images from the Hubble telescope (pictured above). They say the claimed excess was due to inaccurate estimates of emission from zodiacal dust. What's more, they could attribute spatial variations in the background to previously undetected nearby galaxies.

NANOTECHNOLOGY

Spheres inside cells

Environ. Sci. Technol. doi:10.1021/es062541f (2007)

Concerns about nanoparticle toxicity have prompted researchers to look closely at how C_{60} molecules interact with cells.

Alexandra Porter at the University of Cambridge, UK, and her colleagues imaged C_{60} that had infiltrated human macrophages — cells that have a role in clearing debris from the lungs. The researchers showed that a technique known as energy-filtered transmission electron microscopy can pick out the carbon spheres. They could see individual molecules and tell apart aggregates that were crystalline or disordered.

C_{60} appeared in the cells' cytoplasm and nuclei. The molecules were concentrated just inside the cell wall, suggesting that they had infiltrated the cell through its membrane.

GENETICS

Mutations linked to autism

Nature Genet. **39,** 319–328 (2007)
Science doi:10.1126/science.1138659 (2007)

Two studies point to genetic changes that may contribute to the spectrum of autism disorders.

The Autism Genome Project Consortium scanned the genomes of more than 1,000 affected families for single-nucleotide variations in DNA inherited alongside the disorder. They also searched for inherited versions of mutations known as copy-number variants — deletions or duplications of chunks of the genome. Their findings implicate two major regions of DNA, one of which is linked to neuronal proteins called neurexins.

Independently, Jonathan Sebat and Michael Wigler at Cold Spring Harbor Laboratory, New York, and their colleagues compared the role of copy-number variants in sporadic and inherited cases of autism. They found that such mutations appear spontaneously in 10% of patients with sporadic autism, but in only 2% of patients from families with more than one affected member. This suggests that the two classes of autism differ in the primary genetic mechanism involved.

Correction

The Research Highlight 'An added dimension' (*Nature* **446,** 234; 2007) incorrectly referred to *Schizosaccharomyces pombe* as budding yeast. It is fission yeast.

JOURNAL CLUB

James Bauer
College of William and Mary, Gloucester Point, Virginia, USA

A marine scientist marvels at connections between the cold war and slimy mudflat worms.

Having grown up on the coast of New England, my childhood involved a good deal of digging around in the intertidal mud, unearthing things that most people of good sense do their best to avoid — things such as slimy, slithering worms, which often bite or smell bad, or both.

Older but no wiser, I was delighted to come across a recent paper (E. Teuten *et al. Mar. Ecol. Prog. Ser.* **324,** 167–172; 2006) that has cleverly extracted a surprising scientific result from studies of such mudflat worms.

As well as reminding me of my dubious childhood pastime, the work recalls the period in which I grew up, during the cold war, when much of the world lived in fear of the nuclear weapons then being tested. This work takes advantage of one legacy of those tests.

The bomb tests sent into the atmosphere lots of the isotope carbon-14, normally present only at low levels. This bomb carbon-14 subsequently made its way into the oceans, where it became incorporated into plankton. The plankton in turn sank and became part of the coastal mud, providing a home and a food source for marine sedimentary animals.

Mudflat worms are generally believed to ingest wholesale the nondescript sediment in which they live, yet the worms examined in this study contained more bomb carbon-14 than the sediment surrounding them.

Thus, it seems that the worms assimilate from the amorphous goop, material that has been deposited since the cold war and so is younger than the average age of the sediment. Presumably, they do so because the newer material is more nutritious, but how they extract it is unknown.

Makes me want to get back out by the sea with my bucket.

NEWS

Geophysicist faces probe into use of research funds

Anti-fraud officials are investigating one of the founders of a controversial science academy over allegations concerning financial irregularities in European research projects, *Nature* has learned.

Staff at the European Anti-Fraud Office in Brussels have launched an investigation into allegations made against geophysicist Philip Carrion. The office, set up in 1999, has the power to investigate allegations of misuse of European Commission (EC) funds. Spokeswoman Ingeborg Gaspard, confirming that an investigation was ongoing, would not give details of the research involved. But *Nature* has identified three projects involving Carrion, all funded by the EC's Framework programme for research, in which disputes over finances have arisen and complaints been made to the commission. Carrion rejects all allegations of financial impropriety and says he welcomes the investigation as a chance to clear up the disputes.

Carrion first came to the attention of the science community in 2002, after *Nature* investigated the European Academy of Sciences (EAS). The invitation-only academy, which Carrion helped establish in 2002, claimed to have been founded to bring together scientists in Europe and elsewhere, but at the time it seemed to have no publications, projects or meetings, despite listing 600 members (see *Nature* **419,** 865; 2002). Some researchers named as members, including Nobel laureates, said at the time that they had never heard of the organization.

But Carrion and his colleagues have continued to develop the academy, which is based in Liège, Belgium, where Carrion lives. According to the academy's president, Hélène de Rode, a lawyer with Musch, de Pierpont, de Rode, Lahaye in Liège, the academy now has around 650 members. It publishes newsletters and a journal, *Annals*. But Carrion's role as a member of the academy's executive committee ended when he resigned on 16 February. Carrion says he had spent years working for the academy for free and did not want to go on giving his time for nothing. De Rode declined to expand on the reasons for his departure.

Before Carrion's resignation, the academy's website listed him as a professor at the University of Udine, Italy. Gaetano Russo, head of the university's civil-engineering department, says that Carrion was given visiting professor status between 2001 and 2004 in order to work on a project on advanced aluminium alloys. Russo says Carrion agreed to run mechanical tests at Udine that he failed to carry out, and then asked Russo to sign an agreement saying the work had been done and requesting payment of €100,000 (US$130,000) from the EC, to be split between them. Russo declined to sign, and in June 2005 wrote to the commission to make clear that the work had not taken place. For his part, Carrion says he never asked for money and that the Udine lab was not used in the alloys project because it did not have the right equipment.

More recently, disputes have arisen between Carrion and André Leclercq, a human-resources manager with the steel company Arcelor in Liège. Leclercq says he employed Carrion in 2005 as a technical manager on a Framework-funded project on thin-film technology. He says he was introduced to Carrion through the academy.

Carrion says that as part of his role in the Framework project, he agreed to carry out a marketing study and develop a business plan for Api, a business-development organization based in Ancona, Italy, and a partner in the project. He also says he agreed to manage a bank account into which payments for Api's work would be made.

Leclercq says that an advance of €479,000 of EC money was paid in July 2005 into the account controlled by Carrion, in part to pay for the business plan and marketing study. Carrion says he completed the full documents last November, but Leclercq did not receive them until 9 March. By this time a complaint had been made to the commission and a second payment withheld. Carrion adds that he has received no money from the Api account he controls and has still not been paid for his work. Api could not be reached for comment.

A third complaint concerning a project involving Carrion was made in July 2006. Stan Veprek, a materials scientist at the Technical University of Munich, says he contacted the commission after becoming concerned over Carrion's financial management of a project to develop new machine tools. During the project, some funds were transferred into an account named "EAS", but Carrion says these have now been repaid. Carrion adds that a disagreement over payments arose because Veprek did not do the work required. Veprek denies this.

Although no other members of the academy have been linked to the allegations made against Carrion, the organization itself has been the subject of intense criticism from its members. The academy's journal *Annals* is not listed on indexes such as PubMed, and it organizes meetings only infrequently. Boris Verkhovsky, a computer scientist at the New Jersey Institute of Technology in Newark, who has openly clashed with de Rode in the past, speaks for many members when he complains of a lack of transparency about the decisions made by the organization's executive committee.

Investigations are under way into the possible misuse of European Commission money.

C. LAMBIOTTE/EC

De Rode says she has received no official communication from the commission and cannot comment on the investigation or a possible link to Carrion's departure.

On the future of the academy, she says: "There is no link between the well-being and future of the EAS to one of its past executives, Professor Carrion. The same observation applies to Professor Verkhovsky. Both of them were active in the establishment of our academy, but, as time went on, their personal involvement was duly replaced by the activity of committees, which is perfectly normal. The EAS is fulfilling its mission admirably, and the best is still to come." ■

Jim Giles

EMPICS

CHINESE BAN ON INTERNET CAFÉS
If addiction's the problem, is prohibition the answer?
www.nature.com/news/columnsandblogs

Q&A: Klaus Töpfer

Among those paying close attention to the European Union (EU) summit meeting on greenhouse-gas emissions in Brussels last week was Klaus Töpfer. As German minister for the environment and reactor safety from 1987 to 1994 and under-secretary-general of the United Nations and director of the UN Environment Programme in 1998–2006, Töpfer has been keeping an eye on international climate and energy diplomacy for two decades. He talked to **Quirin Schiermeier** about the EU's commitment to cut overall greenhouse-gas emissions by at least 20% by 2020, compared with 1990 levels. These cuts would rise to 30% if the United States and other industrialized countries were to commit themselves to 'comparable' emissions cuts after 2012, and if large developing countries including China contribute 'adequately'.

What do you think of the summit's outcome?
EU leaders know well enough that a 20% reduction is the absolute minimum, that's why they have offered 30% as an additional option. It is the maximum outcome that takes into account the political preconditions in the enlarged EU. In any case, it is a remarkable success for the German Chancellor Angela Merkel.

What more would you have hoped for?
The agreements must be specified further, and, importantly, they must be of a binding nature. We need to go for the 30% target and be aware that the 20% target for renewable energy would benefit from a more structured and more detailed decision on different renewables.

What do you think about the French proposal to treat nuclear energy on equal terms with renewable energies?
Whether or not to use nuclear energy to achieve the target must be the sovereign responsibility of each member state. This is also the EU's position, and it is a very wise decision. [But] increasing nuclear's share in a way that it would make a real contribution to substituting use of fossil energy would vastly increase political, safety and resource issues. We have to do our utmost to invent a nuclear-free energy supply structure.

Do you think the EU agreement will have a positive impact on international climate negotiations for the post-2012 period?
Yes, without a doubt. The leadership the EU has taken gives a clear signal, especially to the fast-growing economies in Asia, that we are aware of the Rio de Janeiro principle of 'common but differentiated responsibility'.

Which concrete measures — technological, political and regulatory — are needed to make sure the EU can achieve the goals agreed on?
With regards to technology, we need a revolution in energy efficiency. Energy efficiency is a chance to achieve a short-term, high return on investments to reduce energy intensity. And this is going over the whole range of energy demand, from buildings to mobility to electrical goods. In political terms, these positive, far-reaching results should pave the way for a Kyoto follow-up beyond 2012. As far as the regulatory framework is concerned, the most important thing is the need to internalize the costs of carbon dioxide in the energy price of fossil fuels. ■

A. BECKSTEIN/REUTERS

Klaus Töpfer is upbeat about the EU's environmental targets.

SPECIAL REPORT

Seed money to bring in pioneers

California's stem-cell initiative has finally handed out its first research grants. But will all the money actually move the field forward? **Erika Check** reports.

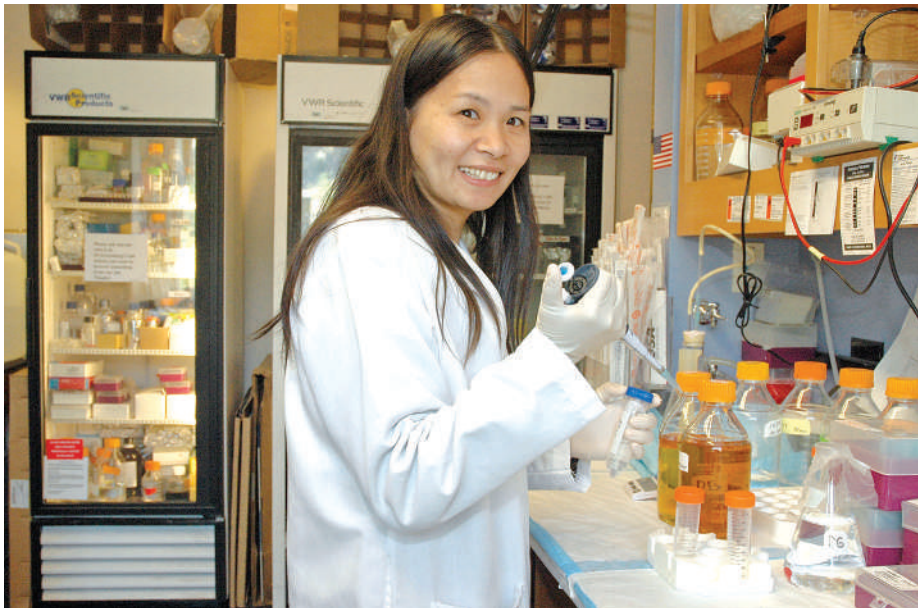
Gregory Kovacs has had no shortage of stimulating work. An engineer at Stanford University who makes biosensors, Kovacs serves as a regular adviser to the US Defense Advanced Research Projects Agency. He helped investigate the 2003 *Columbia* space-shuttle accident. And he climbed the 6,000-metre Licancabur volcano in Chile in the name of medical research — twice.

One thing Kovacs hasn't tried is stem-cell research. But that's about to change. Last month, Kovacs became one of 72 researchers from other fields who have been enticed to work on human embryonic stem cells. All were lured by grants from the California Institute for Regenerative Medicine (CIRM), the state initiative approved by voters in 2004 to overcome federal restrictions on such research. Kovacs, for his part, intends to use his $634,000 grant to study the electrical development of heart stem cells. And he doesn't seem fazed at entering one of the most controversial fields of science, and one in which he has little experience: "At Stanford, you're parking your car next to some Nobel laureate who knows how to work with stem cells," he says. "We can get help."

Kovacs will have even more company on 16 March, when the CIRM hands out further grants worth up to $80 million, this time to established scientists who are already experts in stem-cell research. But it is perhaps more intriguing to look at the initial round of money given to new investigators, the SEED (Scientific Excellence through Exploration and Development) grants. They constitute a rare experiment: to discover whether $45 million and a fresh cast of researchers can affect the future of one of science's hottest fields.

The answer begins with the SEED awardees themselves, who span the gamut from wet-behind-the-ears assistant professors to established scientists with sterling reputations. Twenty-seven of the SEED grants went to scientists who have been working independently for six years or less. Other grants went to highly experienced investigators — even institute directors (see 'Staking it all on stem cells').

The researchers approached the CIRM for a variety of reasons, including the exhilaration of entering a cutting-edge discipline and the opportunity to extend their work into a new system.



R. VENTE

Xianmin Zeng left the National Institutes of Health for stem-cell research at the private Buck Institute.

Another oft-cited factor is the tight budgets at the National Institutes of Health (NIH), the major public funder of US biology research. "In these times of greatly diminished NIH funding, it is fantastic to have a forward-thinking programme such as the CIRM that has stepped up to fill the void," says Warner Greene, a veteran HIV researcher who directs the Gladstone Institute of Virology and Immunology at the University of California, San Francisco.

"Without CIRM funding, I wouldn't have gotten into this, given the barriers involved."

This year, the NIH will fund $37 million in human embryonic stem-cell research, but California researchers such as Greene see their future with the CIRM. Greene received a SEED grant for $778,000 to study how proteins that protect cells against HIV might also help human embryonic stem cells maintain stable genomes.

The typical SEED grant handed out is between $300,000 and $800,000 for two years — roughly comparable to the NIH R01, the typical biologist's bread-and-butter grant. The money from the CIRM is not enough to redirect the entire focus of a laboratory. But it does seem to have been enough to encourage scientists to overcome reservations about a field that is notorious for the difficulties involved, including restrictions on federal funding and the inherent challenges of working with cells that are extremely hard to deal with. "Without CIRM funding, I wouldn't have gotten into this, given the barriers involved," says Anirvan Ghosh of the University of California, San Diego, a leading developmental neurobiologist. Ghosh received $612,000 for work to try to make brain cells from human embryonic stem cells. "The CIRM lowered those barriers enough to give me the motivation to write up a proposal," Ghosh says.

Given the grantees' inexperience with stem cells, others may question how serious they are about staying in the field for the long term. But outgoing CIRM president Zach Hall says the agency's review process weeded out those who were only looking for quick cash. "We had 231 applications, and I'm sure among them there were those whose motivation was primarily to get some money to support their lab," he says. "But the quality of people that we funded is extraordinary."

Now the difficult work begins, as researchers confront the day-to-day realities of working with human embryonic stem cells.

NASA/JPL/SPACE SCI. INST.

LUNAR AND PLANETARY SCIENCE MEETING
Get all the news from the conference in Houston
http://blogs.nature.com/news

First, there's the challenge of just obtaining the cells. Researchers receiving federal funding must, by presidential directive, work only on particular stem-cell lines created before August 2001. Many researchers prefer to work on newer lines, which are seen as more reliable and better characterized, but to do so they must use private sources of funding. Many of the scientists who received SEED grants don't yet have labs funded by non-federal money, and so can work only on the 21 designated lines.

"The first thing we'll do is get the NIH-approved lines, and even that takes time," says David Feldheim, an assistant professor of biology at the University of California, Santa Cruz. Feldheim got $500,000 from the CIRM to study a family of signalling proteins that may affect the development of human embryonic stem cells.

But even those scientists lucky enough to be at a university with a privately funded lab still face issues working with CIRM funding. For instance, Ghosh intends to use his money to study cell lines not approved for NIH support. But it's still unclear whether he can perform part of the analysis using instruments at his university that were built with federal money, such as mass spectrometers. "There are still grey areas, and on the day-to-day issues there is still some hassle," he says.

And although there is plenty of excitement over the first round of SEED grants, others caution against hoping for too much. The patient advocates who helped convince voters to approve the CIRM believe that it will lead to life-saving cures. But veteran researchers have seen high hopes deflated in other fields many times before.

For instance, Stanford University biologist Mark Kay has seen hype lead to disillusionment in the controversial area of gene therapy. Kay received a $640,000 SEED grant to study ways to genetically modify human embryonic stem cells. But he worries that the stem-cell field is even more vulnerable to dashed hopes because of the astronomical expectations that have been set for it, and the way it has become a high-profile political and moral issue.

"Right now, we're in a honeymoon phase," Kay says, "but I'm really worried that this is going to be similar to what happened in gene therapy, where there was way too much hype, things don't come as quickly as people want, and there's a backlash."

For now, such concerns may seem premature. But soon the new grantees will have to start the hard work of actually spending the money. And all eyes will be on them to see whether they can deliver. ■

See also Editorial, page 232.

Staking it all on stem cells

A $730,000 grant would be something of a pittance to one of the University of California schools, which have thousands of faculty and staff and billion-dollar endowments. But that same grant means a lot more to the typical independent research institute with an annual budget of tens of millions of dollars and perhaps as few as 15 faculty members.

That's part of the reason that the Buck Institute for Age Research, a seven-year-old, $30-million independent research institute in Novato, California, has staked its future on the California Institute for Regenerative Medicine (CIRM).

Less than a year after California's voters approved the CIRM, the Buck made two strategic recruitments. It hired stem-cell researcher Xianmin Zeng from the National Institutes of Health (NIH) and chief operating officer James Kovach from the stem-cell company Athersys, based in Cleveland, Ohio.

That expertise has so far helped the Buck win a training grant and a $734,000 SEED (Scientific Excellence through Exploration and Development) grant from the CIRM. The SEED grant will allow Dale Bredesen, the institute's chief executive, to study cell-death pathways in human embryonic stem cells.

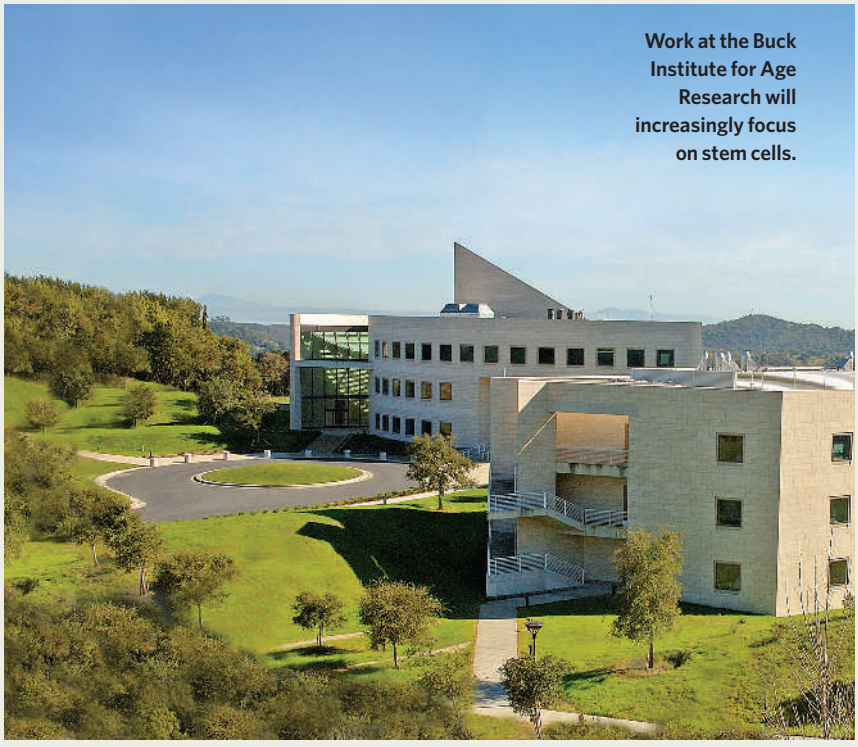
"Support from the CIRM will directly impact on our entire future."

The Buck is in the middle of an expansion, and Kovach says stem-cell research and the CIRM will play an increasingly important role in its plans for growth.

Last October, Kovach spoke to a CIRM working group, and he sums up his message this way: "I said, 'Your support will directly impact on our entire future,' " Kovach says. "Our goal is to partner with other California institutes to be among the first to have our human embryonic stem cells in clinical use."

That may not be an idle dream. Despite their small size, the Buck and California's other independent biomedical institutes have been disproportionately successful in winning CIRM funds so far. Typically, such institutes win 10% of NIH funds nationwide, but in California they have won 30% of the SEED money given out in February.

Kovach argues that the Buck and other institutes made up for their small size with high-quality, innovative CIRM applications. "We won't have as many bullets in the cartridge, but we will have directly interdisciplinary activities that other places don't have," he says. **E.C.**



R. VENTE

Work at the Buck Institute for Age Research will increasingly focus on stem cells.

Microbes reveal extent of biodiversity

Craig Venter is still sailing his *Sorcerer II* sloop around the world, but now he has got 6.3 billion base pairs of microbial DNA to show for it. The pioneer of large-scale genome sequencing has published the first data set from his global quest to sequence microbes from all the oceans.

In a series of papers published this week in *PLoS Biology*, an international team of researchers presents the first results from Venter's Global Ocean Sampling expedition (http://collections.plos.org/plosbiology/gos-2007.php). It follows an earlier, smaller study that uncovered rich microbial diversity in the Sargasso Sea (J. C. Venter *et al*. *Science* **304**, 66–74; 2004). The analyses of the first set of 7.7 million genetic sequences from the expedition reveal that the upper limit on ocean diversity has yet to be set.

"You can over-interpret DNA sequence data, but if you're careful, you just use them as a clue."

The papers point to 1,700 new protein families. Surprisingly, the rate of discovery stayed more or less the same as the number of new sequences grew, suggesting that the number of new protein families will continue to increase. The question, said Venter in a telephone interview from *Sorcerer II* as it bobbed its way through the Sea of Cortez, is how long will that increase continue? And how does one extract meaning from a pile of 6.3 billion As, Ts, Cs and Gs?

The Global Ocean Sampling results reported this week include sequences gathered from 41 locations in 2003 and 2004, from the northwest Atlantic to the eastern tropical Pacific. "The next phase is to try with a much larger data set from the entire circumnavigation," said Venter. "We want to find out if it starts to show any degree of saturation, or is the number of independent protein and gene families so vast that we are still at the earliest stage."

The team has faced some challenges. It has to predict protein sequences on the basis of the DNA sequences it retrieves from its samples. This can be tricky, because not all DNA codes for protein, and some bits of DNA can be read in different ways to produce different protein sequences. It will be important to go back and confirm that the predicted proteins are actually made, says Monica Orellana, a systems biologist at the Institute for Systems Biology in Seattle, Washington. And Brian Palenik, a microbiologist at the Scripps Institution of Oceanography in La Jolla, California, cautions that many sequences that seem to belong to new protein families may in fact be members of known families that contain highly divergent

Bush challenged on funding for children's study

F. CARDOSO/ZEFA/CORBIS

Swings and roundabouts: funding could yet be available to study kids' health.

A controversial US study of environmental effects on children's health got a boost last week, when the congressman who holds the House purse strings for the project vowed to keep it funded despite resistance from the White House.

At a hearing on Capitol Hill on 6 March, Representative David Obey (Democrat, Wisconsin) complained that President George W. Bush's budget request for 2008 eliminated funding for the National Children's Study, as the president had also done the previous year. Authorized by Congress in 2000, the project would follow up 100,000 children from early in the womb until they reached their 21st birthday, examining the effects of everything from video games to chemical pollutants on their health and development. The study is planned for 105 sites nationwide and is expected to cost at least US$3.1 billion over 25 years.

Although the study does not specifically target the benefits of apple pie, it might be expected to enjoy a similar popularity among politicians — who doesn't want to discover the environmental triggers of autism and asthma? But the cost is raising red flags.

"There's no opposition or scientific reluctance on our part" to doing the study, Elias Zerhouni, director of the US National Institutes of Health (NIH) in Bethesda, Maryland, told Obey at the hearing.

VENTER INST.

Sea of change: the *Sorcerer II* is obtaining DNA samples from oceans around the world.

sequences. Still, although the hard numbers might be toned down in the future, Orellana and Palenik don't question that the Global Ocean Sampling database contains a wealth of new proteins.

Already, researchers are tracing the course of protein evolution using the database. Others will search it for new enzymes that might have technological applications, or use it to learn more about microbial ecology. For instance, Jonathan Eisen, a study co-author and microbiologist at the University of California, Davis, will work to match gene fragments in the database with their host organism, a substantial technical challenge when dealing with unfamiliar microbes.

Such studies need to be followed by experiments to establish the function that the gene sequence actually has in the organism, warns Eugene Madsen, a microbiologist at Cornell University in Ithaca, New York. "You can overinterpret DNA sequence data," he says, "but if you're careful, you just use them as a clue and then they either lead to solving the mystery or they don't."

Dennis Hansell, a microbiologist at the University of Miami in Florida, adds that although his research won't entail direct analysis of the Global Ocean Sampling sequence data, the newfound wealth of microbial diversity has made him re-evaluate his picture of how microbes leave their chemical fingerprints in the ocean. He likens the collection of genomic sequences to characterizing all of the pigments an artist could use to paint a portrait. "It's blending the pigments and applying them that results in a Mona Lisa," Hansell says. "I'm looking at the Mona Lisa in my data, and this shows me what the pigments are." ■

Heidi Ledford

See Editorial, page 231.

But Zerhouni said that the agency had to make tough choices to maintain support for bench scientists and, in particular, young investigators. "Clearly, unless there were additional resources, it wouldn't have been wise to sacrifice the next generation [of scientists] for this study," he said.

Obey, who is chair of the House Committee on Appropriations, said that he would solve that problem in 2008 by restoring money for the children's study. "We are going to put that money back next year too," he says. "And it will not squeeze other research because we will expand the institutes' budget, just as we did this year." In February, Obey and his Senate counterparts added an extra $620 million to the NIH's budget for 2007 (see *Nature* **445**, 572–573; 2007), including $58 million to fund the children's study. Early this month, study administrators asked contractors to pitch proposals for managing up to 30 of the 105 sites. For 2008, they are hoping for another $111 million.

> "If there were a cheaper, smaller, faster way to answer the questions, you would do it. But there isn't."

Duane Alexander, director of the National Institute of Child Health and Human Development in Bethesda, Maryland, said that Obey's vow to deliver funding came as "a welcome surprise". The study sets out to answer big questions, Alexander says. "If there were a cheaper, smaller, faster way to answer them, you would do it. But there isn't. And as we get answers about environmental factors, it's going to pay for itself many, many, many times over."

Over 25 years the study is expected to grow up to join the ranks of massive population studies, such as the Framingham Heart Study, that have yielded troves of information on disease causes and associations. If, that is, it doesn't die in its infancy. ■

Meredith Wadman

CHINA DAILY/REUTERS

ZOO NEWS

Vets at an animal research centre in China's Shaanxi province have called for their panda Niu Niu, who lost two-thirds of a leg in a fight last year, to be fitted with a fake limb to allow her to mate without toppling over.

OVERHYPED

Is nothing safe from the perils of global warming? Apparently, the latest victim is the sex trade in Bulgaria, where brothel owners are claiming that their best girls have headed for the hills to entertain snowless skiers.

SCORECARD

▲ **Pluto**
New Mexico has declared that the beleaguered rock should be re-declared as a planet...at least while in the skies over New Mexico.

▼ **Space return flights**
Valentina Tereshkova, the first woman in space, has marked her 70th birthday by declaring herself ready to fly to Mars "without coming back".

ROBOT NEWS

South Korea is doing its bit to prevent future dystopia by drawing up a Robot Ethics Charter, one of the key points of which is that humans should maintain control over our electronic counterparts.

Sources: Reuters, Metro, FPSpace, BBC

SIDELINES

Passive-smoking study faces review

Officials at the University of California are in the throes of a debate on whether to ban research grants from tobacco companies. The discussion has now sparked an independent review of a controversial 2003 report that contested the dangers of second-hand smoke.

Researchers within the nine-campus university can accept money for research from the tobacco industry unless their department specifically bans it. To critics, the new uproar is an example of why all scientists should be prohibited from taking such grants.

"This devastatingly shows university policies and procedures are not robust enough to deal with the influence of tobacco money," says Stanton Glantz, a professor of medicine at the University of California, San Francisco, and a long-time critic of tobacco grants (L. A. Bero *et al. Tob. Control* **14,** 118–126; 2005).

Robert Dynes, president of the university system, has led the fight to continue tobacco funding, saying that stopping it would infringe academic freedom and possibly hinder other business grants. He has also argued that university policies adequately regulate the ethics of such funding.

The new review concerns a study in the *British Medical Journal* (*BMJ*) that said spouses of smokers were no more likely to die of lung cancer and heart disease than were spouses of non-smokers (J. E. Enstrom and G. C. Kabat *Br. Med. J.* **326,** 1057; 2003). The study, led by epidemiologist James Enstrom of the University of California, Los Angeles, looked at 118,000 subjects from a study set up by the American Cancer Society beginning in 1959.

But top scientists at the cancer society say they repeatedly warned Enstrom of possible deficiencies in his analysis — particularly a 25-year gap in which exposure to second-hand smoke could not be verified. The society also says that when it gave Enstrom computerized records of study subjects, it was not aware that he was receiving funding from the tobacco industry. Later tobacco-related lawsuits revealed he had received money from industry funnelled through an organization called the Center for Indoor Air Research. And court records show Enstrom previously did consulting and research for attorneys defending the tobacco companies R. J. Reynolds and Philip Morris.

In August 2006, a US federal judge cited the *BMJ* study as a prime example of how nine tobacco companies engaged in criminal racketeering and fraud to hide the dangers of tobacco smoke. The tobacco companies dispute the judge's decision, which they are appealing.

D. BUTOW/CORBIS SABA

A study that played down the dangers of second-hand smoke is now being scrutinized for bias.

Enstrom and his co-author Geoffrey Kabat, formerly of the State University of New York at Stony Brook, stoutly defend the research against its critics. "They have engaged in a very sophisticated four-year campaign of disinformation about me," says Enstrom. "I have withstood incredible attacks."

The latest round of debate began last autumn when the chief executive of the American Cancer Society, John Seffrin, wrote a letter to the University of California's board of regents arguing that tobacco funding should be banned. In the 12 October letter, Seffrin argued that tobacco-funded front groups "publicized misleading results" while giving "the false implication" that the society had endorsed the study. He cited Enstrom's *BMJ* article in particular, alleging that Enstrom "ignored" complaints of "fundamental methodological problems".

> "This shows university policies are not robust enough to deal with the influence of tobacco money."

But the letter never made it to the regents at the time, because of what Dynes now calls an error in handling correspondence; he has initiated a review of such practices. The letter became more widely known when Glantz showed it on an overhead screen during a regents meeting on 18 January.

The next day, Wyatt Hume, provost at the University of California's president's office, wrote to Seffrin saying that the university "takes allegations of scientific misconduct extremely seriously". If there is "specific information in support of an allegation of scientific misconduct against Enstrom", he wrote, he would relay it to officials at the Los Angeles campus so that they "can pursue the matter further". Shortly after, officials at the cancer society sent a seven-page list of what they cited as issues with the *BMJ* article.

On 16 February, Dynes wrote a four-page letter to regents chairman Richard Blum noting how the allegations about the *BMJ* article didn't follow the form of a typical complaint. "We decided to err on the side of caution and refer the matter to the appropriate campus official for follow-up," Dynes wrote. Officials at the Los Angeles campus "will conduct a thorough review of the documents", he wrote, and "will take further steps to determine whether any research misconduct took place".

In an interview, Enstrom acknowledged receiving the various letters and corresponding with the University of California's authorities. "I am working on this with regents' approval," he said. "I am being allowed to defend myself by the appropriate people."

He "absolutely" denies any misconduct in the study. And Kabat objects to the university's regent policies being based "on allegations motivated by a political agenda and unsupported by any facts".

The university's regents are expected to revisit the debate on tobacco funding in May. ■

Rex Dalton

Blakemore steps down from Medical Research Council

Neuroscientist Colin Blakemore is to give up his position as head of Britain's Medical Research Council (MRC) when his present contract expires this September.

Blakemore, who has been in post since 2003, told *Nature* that he has decided not to apply for a second term because the MRC will suffer a temporary loss of independence while new advisory boards are established to oversee UK health funding. The MRC distributes around £500 million (US$960 million) of government money per year.

He departs with support from many of the researchers he funds. Over-optimistic spending had left the council short of money before Blakemore's arrival; he stabilized the funding situation and simplified grant procedures.

Sacked professor launches legal action to regain job

A professor dismissed by the University of Tokyo in Japan said on 2 March that he will take the university to court to demand his job back.

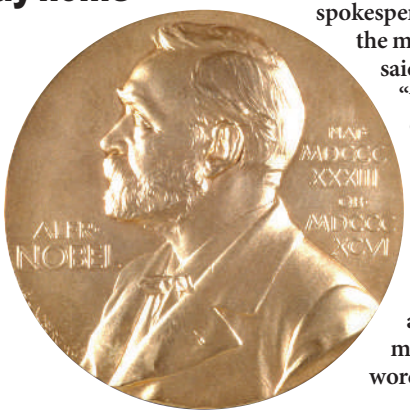
Kazunari Taira was fired in December after an investigation found "no reproducibility and no credibility" in four RNA papers from his lab. Also sacked was the researcher who had been in charge of the questioned experiments (see *Nature* 445, 12; 2007). Both deny involvement in misconduct.

The university criticized Taira for "neglect on appropriate supervision and lab management". But Taira argues that the case could put researchers off running labs and wants the university to restore his professorship and resume paying his salary. A spokesperson for the university says that it is considering its response to the lawsuit. Hearings for a provisional ruling will begin this month.

Nobel medal stolen from cabinet finds its way home

A Nobel medal swiped from the University of California, Berkeley, was returned to officials in a brief ceremony on 7 March.

The solid gold medallion, awarded to physicist Ernest Lawrence in 1939 for the invention of the cyclotron, was stolen from its display case at the end of February. It has an estimated monetary value of $4,200.



NOBEL FOUNDATION

Authorities last week charged a biology student at the Berkeley campus with the theft. The 22-year-old worked at the university's Lawrence Hall of Science, where the medal was stored. He told police that he used a key to open the case, taking the 200-gram medal on a whim. Plans are now under way to house the medal in a more secure manner.

Wildlife agency accused of gagging scientists

The US Fish and Wildlife Service is fending off accusations of scientific censorship over a memo that outlines restrictions on who can say what about climate change, sea ice and polar bears.

The memo states that agency personnel travelling in northern countries should indicate a spokesperson and provide assurances that they "understand the Administration's position on these issues". A travel request for a non-spokesperson, included in the memo as an example, said that the individual "will not be speaking on or responding to these issues".

Hugh Vickery, a spokesman at the Department of the Interior, the service's parent agency, says the memo was "badly worded" and really only applied to discussions of policy. "It's a reminder not to get into areas where they shouldn't go." The polar bear has become political because is it being considered for a slot on the US endangered species list, thanks to the threat of melting sea ice.

Scientists plan defence against asteroids

There's no doubt that a space rock slamming into Earth could cause substantial damage, but exactly what humans should do about the threat has not yet been decided. That's why scientists gathered for the Planetary Defense Conference in Washington DC on 5–8 March. Their aim was to compose a white paper on the subject — the first to be mandated by the US Congress.

Scientists at the conference said that it would cost about US$1 billion to find at least 90% of the 20,000 estimated potential Earth-killers by 2020, and discussed how a space rock on a collision course might be deflected. Options range from using spacecraft as 'tugboats' to drag an object into a new orbit, to proposals that rely on nuclear detonations to knock a rock off target — similar to the strategy used against a comet in the 1998 film *Deep Impact* (pictured). The white paper will be published at www.aero.org/conferences/planetarydefense.

DREAMWORKS/PARAMOUNT/KOBAL COLLECTION

The diaries of Darwin's wife debut online

For a unique glimpse into the family life of Charles Darwin, historians can now delve online into the diaries of his wife Emma.

Emma noted in pocket diaries details of her everyday life such as shopping costs, dinner parties and the illnesses of her children and husband. Spanning from 1824 — before the pair's marriage in 1839 — to 1896, the diaries belong to Darwin's great grandson and are looked after by the Darwin Archive at the University of Cambridge, UK. Images of the diaries' pages were made available to browse at http://darwin-online.org.uk/EmmaDiaries.html on 6 March.

The diaries record the visits of friends and scientists to the couple's house. For example, an entry in April 1856 says "Lyells came". A note in Darwin's own journal, dated a few weeks later, tells what happened after the visit by Charles Lyell and his wife: "Began by Lyells advice writing species sketch."

Correction

The News Feature 'The new face of the Arctic' (*Nature* **446,** 133–135; 2007) mistakenly located the community of Tuktoyaktut in Alaska. It is in Canada's Northwest Territories.

Washington, we have a problem

Does anybody care about NASA? Most in Congress do only if there is a research centre in their district. **David Goldston** explains why members never step up to the plate to set priorities.

Congress is facing some tough issues as it begins to figure out NASA's budget for the next fiscal year. A new report from the National Academy of Sciences decries the state of Earth sciences at the US space agency; NASA claims that congressional trims to its current budget for human spaceflight will delay the debut of its Orion spaceship; and the Bush administration is proposing deep cuts in the agency's aeronautics budget for next year. What is Congress likely to do?

Probably the usual — just kick the can down the road a bit. Congress often alters the proposed NASA budget, but it rarely, if ever, manages to make fundamental decisions about what the agency's priorities should be. That's why NASA so often looks like it is carrying a broad portfolio of programmes, all of them seeming as though they are on life support.

There are basic institutional reasons for this ongoing paralysis, and they're unlikely to change any time soon. Very few in Congress pay much attention to NASA, and those who do tend to have fierce loyalties to just one aspect of the agency's activities. As a result, most efforts to set clear priorities end in stalemate.

The most active members of Congress generally represent NASA facilities, and they work to protect their centre's budget and jobs. To take some obvious examples, Senator Kay Bailey Hutchison (Republican, Texas) throws her weight behind the human spaceflight programme to ensure that Houston doesn't have a problem. And Senator Barbara Mikulski (Democrat, Maryland) watches out for her state's Goddard Space Flight Center.

That's not to say that policy plays no role in debates. Congress has shown strong support for NASA's aeronautics programmes because aviation is seen as critical to US competitiveness and because it has been a key NASA responsibility from its inception in 1958. Led by members representing NASA's aviation centres, such as Representative Frank Wolf (Republican, Virginia), Congress has often rejected steep cuts in aeronautics even though the aviation industry has barely lifted a finger on behalf of the programmes. (Industry may be reluctant to alienate NASA leadership over aeronautics when it has far more money riding on human spaceflight programmes, and NASA is shifting its emphasis away from the industry favourites, demonstration programmes.) The result is that aeronautics funding has declined, but at a slower rate than proposed, leaving NASA in a kind of limbo: aeronautics lacks the money for its programmes to thrive, but it still consumes dollars that NASA's leaders want to shift into higher-priority areas.

PARTY OF ONE

Note that none of these NASA issues breaks along partisan lines. President George W. Bush's lunar mission raised howls equally from liberal Democrats, who thought the money ought to be spent on Earth, and conservative Republicans, who thought that government ought not spend so much money, period. But the whole matter was quickly forgotten by most.

So the story one hears now from most members of Congress, and some in the media, is that the president made a speech about going to Mars in 2004, got nothing but grief for it, and the proposal went nowhere. This is, of course, almost entirely wrong. The bottom line of the president's speech was to return astronauts to the Moon by 2020. That programme has been going forward steadily (see *Nature* **445,** 474–478; 2007), albeit with less funding than originally proposed. The president's silence has been, if anything, a strategic retreat that has actually hastened plans for a lunar mission — because few other than NASA's most ardent supporters in Congress are paying much attention.

Congress had a chance to debate NASA's future in 2005, when it considered a bill to reauthorize the agency. Authorization bills don't provide funding; they set policies and lay out Congress's multi-year intentions for an agency. The bill dealt only with NASA, unlike spending bills where NASA money is part of a package that includes funding for unrelated agencies.

The bill, as usual, reflected the congressional stalemate. It both endorsed the lunar mission and explicitly required NASA to continue funding aeronautics, and Earth and space science. It indicated that more money than the president was proposing would be needed to continue all those activities effectively, and it implicitly endorsed the president's plan to retire the space shuttle in 2010 — which Texas and Florida representatives had considered blocking. (As chief of staff of the House Committee on Science, I helped write and move the bill.)

No bill that made tougher choices had a prayer of making it through the committee stage in either the House or Senate. And in any event, Tom DeLay (Republican, Texas), who was then House majority leader, would have prevented any bill that failed to endorse the lunar mission from coming before the House. When the bill did come up, there was virtually no debate. The only member in the House who spoke against it was a liberal Democrat, Barney Frank (Massachusetts), who said the human "mission to Mars" was a waste of money.

The inattention of Congress mirrors that of the public. Members of Congress hear from very few constituents about NASA unless they represent NASA employees. Polls, depending on how they are worded, can show strong public support for human spaceflight or little, which was true even in the Apollo era.

The one time in recent memory that Congress heard something of a public response to a NASA decision was when the agency announced (or rather, let seep out) that it was cancelling the mission to extend the life of the Hubble Space Telescope. I even heard a waiter in the congressional dining room tell then-NASA administrator Sean O'Keefe to "save that Hubble".

As the public objected, Congress held hearings and made noises about reversing the decision. But it was in a tough position because O'Keefe had claimed that the decision was meant, in the wake of the *Columbia* disaster, to reduce the risk to shuttle astronauts visiting Hubble. Congress was also reluctant to override what was an operational, as well as a policy decision. So paralysis again seemed the order of the day. It took the new NASA administrator, Michael Griffin —not Congress — to break the logjam by reversing the decision altogether. ■

David Goldston is a visiting lecturer at Princeton University's Woodrow Wilson School of Public and International Affairs.

BUSINESS

Quantum leap of faith

A Canadian company says it is the first to bring a quantum computer to market but, as **Geoff Brumfiel** reports, not everyone is buying into the approach.

The public debut of what is being billed as the world's first commercial quantum computer was a low-key affair. And so it should have been, say critics.

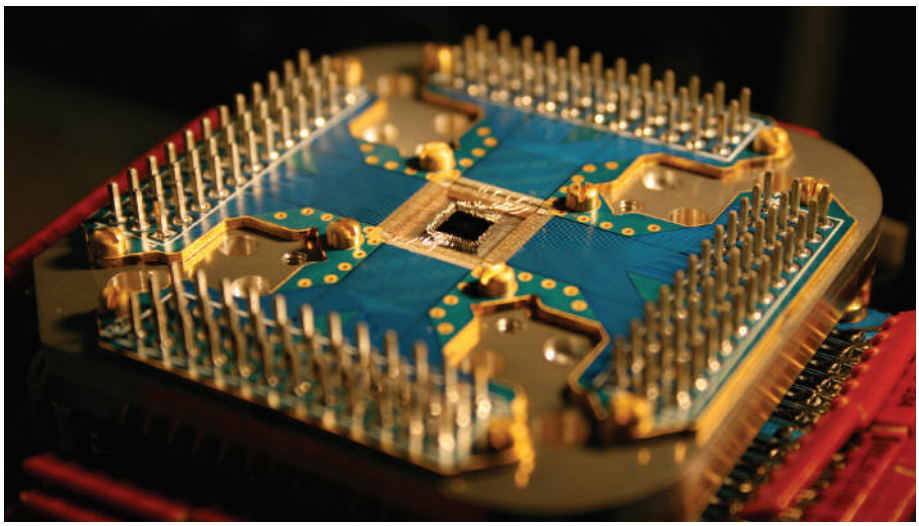
D-Wave Systems, a private company based in British Columbia, Canada, launched its machine last month at the Computer History Museum in Mountain View, California. Geordie Rose, D-Wave's founder and chief technology officer, tried to persuade an audience of scientists, journalists and entrepreneurs that the machine — called Orion — could be used to solve awkward problems such as how to seat a wedding party to minimize friction between the guests.

"We want to build a machine that cannot be bettered by any machine that obeys the laws of physics," says Rose, whose company has adopted a rapid-fire approach to product development. "The philosophy behind our design is build the thing as fast, as dirty and as quickly as you can."

Quantum specialists greeted the debut with scepticism, however. In the run-up to the announcement, D-Wave failed to convince outside researchers that its machine actually works in a quantum regime, says Scott Aaronson, a computer scientist at the University of Waterloo in Ontario, Canada, who writes programs for quantum computers. And even if it does, he says, its 16 quantum bits (qubits) are insufficient to outperform conventional computers: "Even if it worked perfectly, it's still completely useless from an industrial perspective."

Others are a little more sanguine. The approach of going ahead and building the machine without much in the way of proof that it will work is risky, but might bear fruit, says Seth Lloyd, a mechanical engineer at the Massachusetts Institute of Technology in Cambridge who specializes in quantum computing.

Where conventional computers rely on electrons to encode 'bits' at one or zero, quantum computers encode their information in the different quantum states of electrons, photons or atomic nuclei. Relationships between these states are known as 'superposition' and 'entanglement' in the quantum world, and theorists think that they could allow the machines to crunch numbers much more quickly than their conventional counterparts. In particular, they should be able to identify the factors of large numbers much faster, making them valuable for code-breaking. They might also be able to solve a simple problem from a vast set of possibilities, as the example of the wedding seating plan sought to demonstrate.



J. CHUNG/D-WAVE SYSTEMS

D-Wave's 16-qubit processor uses an untested approach to solve complex problems.

Major computing firms such as Hewlett-Packard and IBM have small research groups working away at quantum computation, but see it as a long-term prospect. The main challenge is to build a machine that can set, maintain and read quantum states without being overwhelmed by surrounding noise, specialists say. Control systems need to be exquisitely sensitive, and the very act of writing or reading can disrupt their operation. Most existing quantum computers look like physics experiments, and contain very few actual qubits.

> **"This project has never been about science, it's about building a machine."**
> — Geordie Rose

D-Wave's approach is known as adiabatic quantum computing. It uses a superconducting microchip to create spinning loops of current, which it says act as qubits, spinning in both directions at once. The qubits then entangle together, putting the entire machine in a 'quantum ground state'. By slowly changing the magnetic fields around the qubits, the computer's operator can then change the computer's entanglement, moving it into a new ground state. Reading this state can, according to the designers, reveal a solution to a complex problem without the need to directly read the state of individual qubits.

Aaronson says, however, that external noise could still be a problem for the machine. He also notes that the company has provided no public evidence that it has its system under control — or that its qubits are entangled correctly. According to Lloyd, researchers are still debating whether the entire approach is of any use: "Nobody knows exactly whether adiabatic quantum computing will work or not."

No rush to publish

But Rose says that the company's scientists are satisfied that the system is indeed operating properly — and that the objective is to try something out and see whether it works, rather than to pin down exactly how it does so. "I understand the scientific community would like to see publications coming from us," he says. "But this project has never been about science, it's about building a machine."

And D-Wave has managed to persuade key investors to stick with the concept, raising some US$38 million since it was founded in 1999, half from the Canadian government and half from venture-capital firms, including Draper Fisher Jurvetson of Menlo Park, California. But even Rose still sees success as an outside bet, and Lloyd shares his reservations. "Investing in something like D-Wave is only for someone who wants to take risks," Lloyd says. "Big risks." ■

We are family

Updating the tree of life needs both the skills of evolutionary biologists and the data from genome-crunchers — the two ignore each other at their peril. **John Whitfield** reports.

On 1 July 1858, in the Linnean Society of London's imposing neoclassical building on Piccadilly, biology changed for ever. That evening John Bennett, the society's secretary, read out papers by two biologists, Alfred Russel Wallace and Charles Darwin. From that point on, linnaean science was about more than describing and classifying living things. It was about using their traits to reveal their evolutionary relationships and assign them a position on life's family tree. Famously, such a tree was the only illustration in Darwin's *On the Origin of Species*, published the following year.

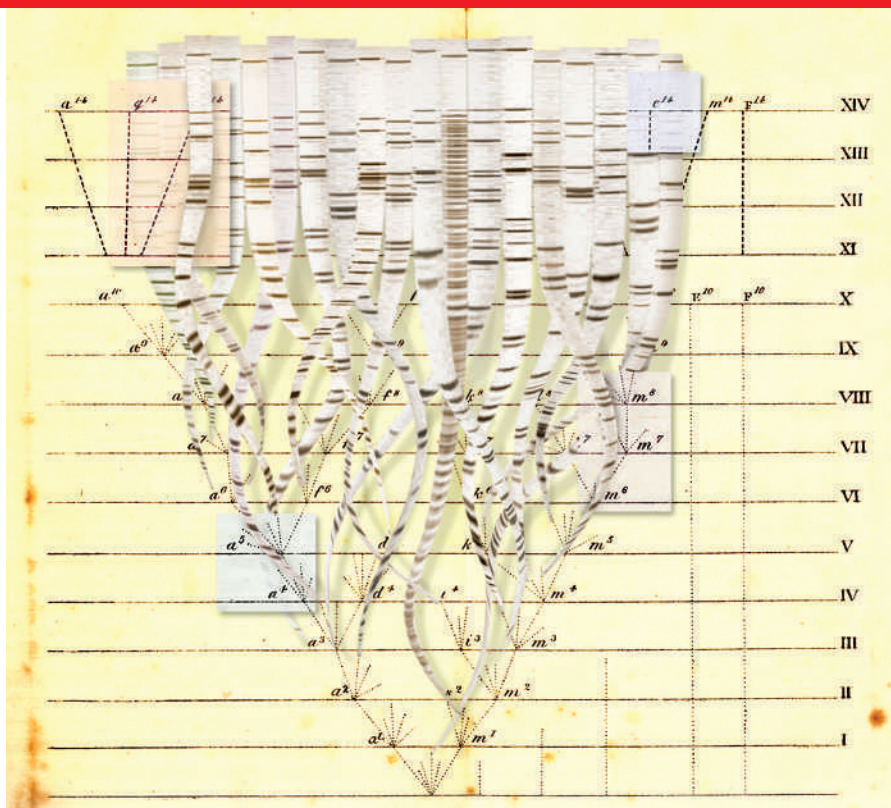
In the past 30 years, DNA sequencing has revolutionized this project. Comparing gene sequences has revealed not only a new domain of life, the archaea, but has affirmed that chimpanzees rather than gorillas are our closest relatives. And what was a trickle has become a flood — the major sequencing centres alone generate some 5,000 bases of DNA a second. About 1,000 species have now had their genomes sequenced, with more being published each month.

This torrent is set to reshape systematics — as the study of evolutionary relationships is known — once more. "Genomes are giving a much better view of the tree of life on Earth," says evolutionary biologist Mark Blaxter of the University of Edinburgh, UK. "The revolution is just starting — every new genome is causing rethinking." Blaxter is one of several researchers who has rallied under the banner of phylogenomics — the use of large quantities of genetic data (not just entire genomes) to build evolutionary trees, or phylogenies.

What has also become clear is that many problems cannot simply be battered into submission with more data. "Questions that are not resolved by a kilobase of sequence are seldom resolved by a megabase," says Jeffrey Boore of the Joint Genome Institute in Walnut Creek, California. At its best, phylogenomics is a two-way street — genome researchers need evolutionary biologists to help them work out the function of genes, and to avoid embarrassing mistakes.

Perhaps the most high-profile gaffe was the declaration by the Human Genome Project in 2001 that 100–200 genes in humans had come directly from bacteria. Analysis had revealed genes found in both humans and bacteria, but not in any species more closely related to humans. Project scientists concluded that such genes probably got into humans by lateral gene transfer from bacteria, a kind of inter-species sex where chunks of DNA cross from one cell to another. The result was heralded as one of the project's major revelations.

Jonathan Eisen first heard this finding while watching the press conference on TV at his workplace, the Institute for Genomic Research in Rockville, Maryland. "I felt sick to my stomach," he says. "They were talking about genes involved in brain development having come directly from bacteria. If you just think about that for a minute, it sounds so implausible." Sure enough, Eisen, and other similarly flabbergasted evolutionary biologists, rapidly shot the claim down, showing that the genes in question were more likely to have been present in the common ancestor of humans and bacteria but then lost in other lineages[1].



Updating Darwin: how do sequence data affect the family tree?

N. SPENCER/TEK IMAGE/SPL

This story is not the only example of the mistakes that can happen when genomics ignores evolution. "Molecular biologists are prone to not treating evolution as a particularly rigorous science," says Eisen, who now works at the University of California, Davis. "But you can't do good genome analysis without evolutionary analysis." Phylogenomics is about integrating the two, he says.

Form and function

Despite the shaky start, genome-sequencing centres now recognize the importance of evolution, and Eisen is one of several evolutionary biologists working at the heart of genomics. It's not just about keeping egg off faces. The techniques of evolutionary biology can also enhance other aspects of genome analysis — finding genes, and working out what is done by the proteins they encode.

Presented with a gene of unknown function, scientists can search for similar genes, of known function, in other sequences. A more rigorous approach is to use the sequence data to build a phylogeny of the different versions of that gene. This lets researchers track the gene's evolution, see how its function might have changed, and identify which other gene is most closely related to the target — which is not necessarily the one with the most similar sequence. Software packages can automate this process, and will predict protein function more accurately than doing a simple sequence comparison.

Tree-building is still the trickiest part. Systematists work out the relationships between genes, species or higher groups by searching for characters, be they morphological, biochemical or genetic, that arose in their last common ancestor and are shared by their descendants — lactation in mammals, for example. But even for just 10 species, there are more than 34 million possible evolutionary trees. Phylogenetic researchers

"The evolution of the animals has plagued us for years. We're hitting a wall." — Antonis Rokas

C. KENNEDY/KRT/NEWSCOM

Family strife: humans, fruitflies and nematodes were the first animals to have their complete genome sequenced — yet their exact evolutionary relationship still remains obscure.

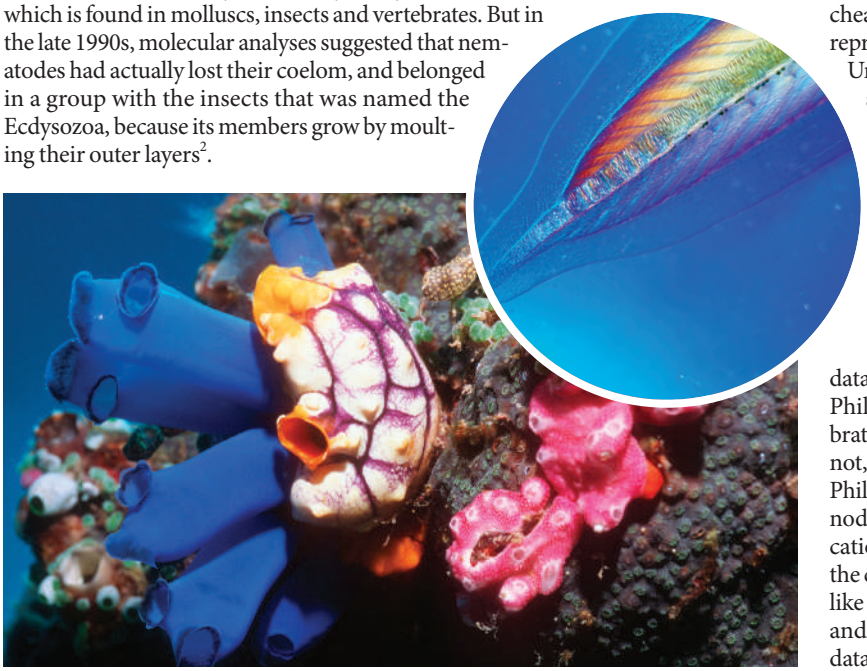
have developed complex algorithms to search among these possibilities, to find the tree most likely to reflect reality.

Picking the right tree is challenging. "For 15 years, people hoped for some software that could solve these problems at the push of a button, but sadly that's not come to pass. It's a very, very difficult problem algorithmically," says Boore. It is especially hard to say whether traits were inherited from a common ancestor, whether they look similar but evolved independently in unrelated groups — such as the wings of birds and bats — or whether they might have evolved but were then lost in later descendants, such as eyes in cave-dwelling fish. Often several different trees are equally good explanations of the data. And different genes from the same set of organisms often predict different trees.

Branching out

It is in the latter case that genome-sized datasets should be helpful. Analysing many genes at once allows the overall pattern of evolution to emerge, swamping the effect of quirky genes and confusing traits. One early success was in working out the relationships between groups of mammals, such as rodents, bats, carnivores and ungulates. These seem to have evolved very rapidly after the extinction of the dinosaurs, so their various common ancestors had little time to evolve unique characters that would mark their descendants. But multiple analyses, combining a few dozen genes from a cell's nucleus with the complete genomes of mitochondria, a cellular energy generator with its own small chromosome, have resulted in a robust consensus on who is related to whom.

Not everything is so tidy. Take the nematodes. For the past decade, a debate has swung back and forth about where this group of worms belongs. Anatomically simple organisms such as nematodes are a particular headache for systematists, as they offer few clues to their ancestral state. Based on morphology, nematodes were placed close to the root of the tree of animals, because they lack a body cavity called a coelom, which is found in molluscs, insects and vertebrates. But in the late 1990s, molecular analyses suggested that nematodes had actually lost their coelom, and belonged in a group with the insects that was named the Ecdysozoa, because its members grow by moulting their outer layers[2].



G. DOUWMA/SPL

Sea squirts briefly usurped cephalochordates (inset) as the closest relatives of vertebrates.

Where to put nematodes, and how they relate to insects, matters to genomics more broadly because the worm *Caenorhabditis elegans* and the fruitfly *Drosophila melanogaster* are two of the most important model animals. If the Ecdysozoa is the true group, both are equally distant from humans. If not, the fly is more like us than the worm. Humans, *C. elegans* and *D. melanogaster* were among the first animals to have their genomes sequenced. These complete genomes seemed to tell a different story from the earlier genetic analysis that put forward the Ecdysozoa. Or rather, they retold the old story. A tree built from the human, fly and worm genomes[3] makes the nematodes the outsiders.

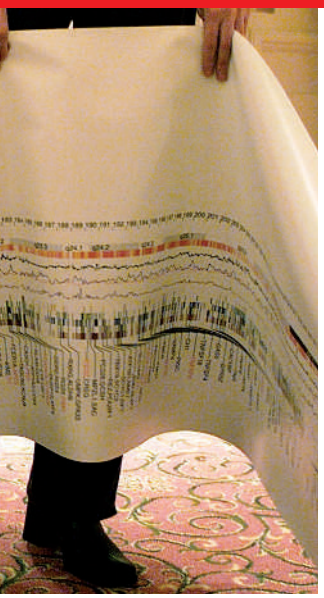
Express yourself

One way to choose between contradictory trees is to add more species to the analysis. With animals, one soon runs out of sequenced genomes to add — but there is a third way between having some data for lots of species and lots of data for a few species. This is to use expressed sequence tags (ESTs), which are DNA sequences made from the genes that are active in cells. EST libraries can provide data on several hundred genes for a species relatively quickly and cheaply, allowing more species to be analysed. Using ESTs representing nearly 150 genes, Hervé Philippe's team at the University of Montreal, Canada, built a tree containing 35 species that supported the existence of the Ecdysozoa[4].

C. NURIDSANY & M. PERENNOU/SPL

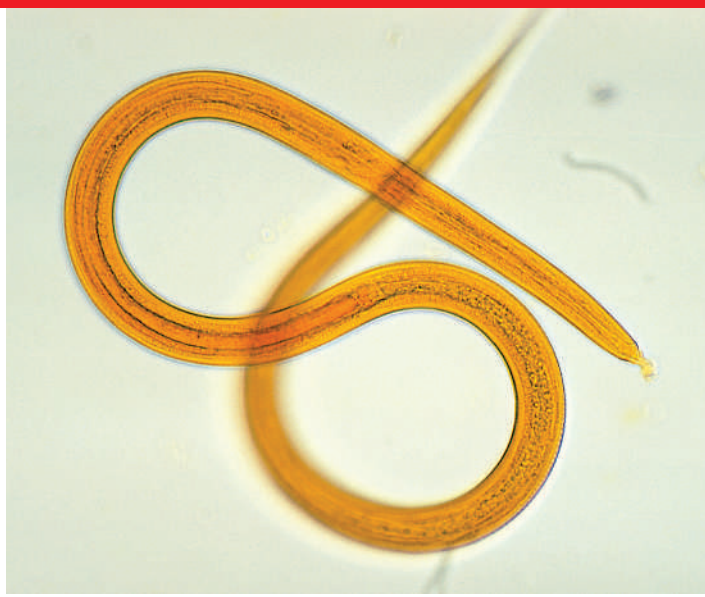
Most phylogenomics researchers are now inclined to believe the Ecdysozoa data. But enough analyses suggest the contrary for the question to remain open, says Antonis Rokas, a genomics researcher at the Massachusetts Institute of Technology in Cambridge. "I would be agnostic," he says. "Just looking at the phylogenetic evidence, I'm not convinced that either side has it right."

This seesawing has become common, as accumulating data tip an argument first one way, then the other. Last year, Philippe's team suggested that the closest relatives of vertebrates were the sea squirts, which resemble bags of jelly, and not, as previously thought, the fish-like cephalochordates[5]. Philippe's work grouped the cephalochordates with the echinoderms (sea urchins and the like) with the troubling implication that either fish-like creatures evolved twice, or that the common ancestor of vertebrates and sea urchins looked like a fish. In November, evolutionary biologist Max Telford and his colleagues at University College London, added a dataset of 35,000 amino acids from other groups, including starfish. The new tree put the cephalochordates back where



GRAPHIC SCIENCE/ALAMY

L. TELFORD



M. MELVIN/CDC

they started[6]. "There was a big sigh of relief," says Telford. "Not even Philippe had much faith in the first result."

In general, relations between the animal phyla — large groups such as molluscs or arthropods — remain a mystery. "The evolution of the animals has plagued us for years," says Rokas. "We're hitting a wall." The animal phyla all seem to have evolved very rapidly some 600 million years ago and the signal in DNA data has eroded. Rokas has compared evolutionary trees of fungi and animals built using the same genes[7]. Those from fungi, whose genes are thought to have changed at a steady pace, are well resolved. The animal trees are rickety.

Will more sequences help? "There are some problems that you could throw a whole genome at and will still probably be unresolved," says Telford. He gives the example of predatory marine worms called the chaetognaths, or arrow worms, which refused to reveal their evolutionary allegiance even when more than 70 genes were analysed[8]. "We had more data than you can shake a stick at, and we still weren't able to place them," he says.

Location, location, location

Where sequence fails, other features of the genome might help. The positions of genes, for example, are also inherited, and changes in gene order can mark evolutionary splits. Boore pioneered this approach, using the order of genes on mitochondria to show that insects and crustaceans are closer to each other than either is to spiders or centipedes[9].

Similarly, the points at which the 'jumping genes' called transposons insert themselves into sequences have been used to reconstruct the history of the mammals[10], and the positions of introns — gene pieces that are cut out before DNA is turned into protein — have been used to support the existence of the Ecdysozoa[11]. When a rare genetic event is shared among a group, it's a sure sign that they had a common ancestor, says Boore.

The usefulness of such approaches is not yet universally accepted. "I have found many genome features don't work well for evolutionary reconstruction — they're too prone to convergent evolution," says Eisen, explaining that the same gene order probably often arises independently in different groups. Philippe believes that sequence data are still the most reliable, but that other approaches will mature in time.

Part of the problem may be that most of the genomes sequenced so far, besides humans, have been lab models, diseases or economically important organisms. None is ideal for understanding evolution — model species and crops have quick generation times, so they are among the fastest evolving species. "The organisms chosen as models have got biological properties that mean, genomically, they're odd," says Blaxter.

"You could throw a whole genome at some problems and they will still probably be unresolved." — Max Telford

But evolutionary biologists are hoping that genome sequencers will soon switch focus. "A lot of genome projects are finishing and the centres are looking for something to do," says Tim Littlewood of the Natural History Museum in London. One target might be the eukaryote tree. The eukaryote domain covers organisms whose cells are divided into compartments, including such familiar kingdoms as animals, plants and fungi, but also a plethora of single-celled organisms called protists. No one has any idea which of these groups is the closest ancestor to the protist that made the leap to multicellularity and gave rise to the animals. "One of the most interesting questions that phylogenomics can address is of the major lines of eukaryotic evolution," says Philippe.

But evolutionary puzzles are not going to drive sequencing agendas on their own, says Boore. One fear is that once all the crops and pathogens are sequenced, the cash will dry up. "So far, the appetite for comparative genomics has been surprising," Boore says, "but it's hard to predict how things are going to go." When selecting organisms for sequencing, the potential for illuminating human biology and disease are strong arguments. Fussing over the lineage of obscure marine species is less compelling.

"Genomics people recognize the need for systematics," argues Littlewood. If systematists — whose discipline labours under a fusty, arcane image — show that voguish molecular research depends on their work, it should help reverse the decline in jobs and funding for the discipline, he says. "In many ways, genomics stands to learn more from systematics than vice versa — the onus is on systematists to point out that we can help." ■

John Whitfield is a science writer based in London.

1. Genereux, D. P. & Logsdon Jr, J. M. *Trends Genet.* **19,** 191–195 (2003).
2. Aguinaldo, A. M. A. *et al. Nature* **387,** 489–493 (1997)
3. Blair, J. E., Ikeo, K., Gojobori, T. & Hedges, S. B. *BMC Evol. Biol.* **2,** 7 (2002).
4. Philippe, H., Lartillot, N. & Brinkmann, H. *Mol. Biol. Evol.* **22,** 1246–1253 (2005).
5. Delsuc, F., Brinkmann, H., Chourrout, D. & Philippe, H. *Nature* **439,** 965–968 (2006).
6. Bourlat, S. J. *et al. Nature* **444,** 85–88 (2006).
7. Rokas, A., Krüger, D. & Carroll, S. B. *Science* **310,** 1933–1938 (2005).
8. Matus, D. Q. *et al. Curr. Biol.* **16,** R575–R576 (2006).
9. Boore, J. L., Collins, T. M., Stanton, D., Daehler, L L. & Brown, W. M. *Nature* **376,** 163–165 (1995).
10. Nishihara, H., Hasegawa, M. & Okada, N. *Proc. Natl Acad. Sci. USA* **103,** 9929–9934 (2006).
11. Roy, S. W. & Gilbert, W. *Proc. Natl Acad. Sci. USA* **102,** 4403–4408 (2005).

The species and the specious

For some, species are simply the things you save; but for taxonomists, the concept is much more complex. **Emma Marris** asks whether Linnaeus's legacy is cut out for conservation.

Conservationists around the world were delighted to hear that the United States was considering adding polar bears to the list of animals enjoying the protection of the Endangered Species Act — the cornerstone of US environmental law since 1973. But some experts may have greeted the news with a wry smile. Polar bears might well be threatened by the rapidly changing climate of the Arctic — but whether they actually constitute a species is up for debate.

Genetic studies have shown that some brown bears (*Ursus arctos*) are more closely related to polar bears (*U. maritimus*) than they are to some other brown bears[1]. According to some interpretations of the word species, this means that if brown bears are a species, then polar bears are not. This is not, in itself, an insurmountable obstacle to protection under the act, which can be quite flexible on taxonomic matters. But negotiating that flexibility is far from easy: when modern taxonomy comes up against a conservation agenda, things can get very complex. Not everything that humans want to save is covered by an easy definition, whether it is in terms of genes, anatomy or ecological role.

Take a less iconic mammal that has been in the act's purview rather longer: Preble's meadow jumping mouse (*Zapus hudsonius preblei*). It's a threatened tea-cup-sized rodent with comically large feet and a counterbalancing tail, and spends its life hopping about the foothills of the Front Range in Colorado and Wyoming. Or that's how some would have it. Others say that the subspecies, named after the naturalist Edward Alexander Preble, is a spurious one, and that the creatures called by that name are just plain old meadow jumping mice (*Zapus hudsonius*), a species with which the United States is crawling.

To protect and serve

The disagreement is not just a matter of status games between mice. The streamside habitat preferred by the foothill mice is also prime real-estate land for residential and commercial development. The degree of protection accorded to the mice thus has implications for developers.

After petitions were filed by a group calling itself the Coloradans for Water Conservation and Development and the State of Wyoming's Office of the Governor in 2005, the US Fish and Wildlife Service announced that it planned to drop Preble's mouse from its list of threatened species and subspecies because of new genetics work by a group led by Rob Roy Ramey when he was at the Denver Museum of Nature and Science[2]. Papers and rhetoric began to fly. A group led by Tim King from the US Geological Survey in Kearneysville, West Virginia, was commissioned by the Fish and Wildlife

VISUAL&WRITTEN SL/ALAMY

A proposal to list polar bears as protected avoids the issue of whether they are a species.

Service to do further genetic tests. It disagreed with Ramey and his group[3]. In the end, the agency called in an outside group — the Sustainable Ecosystems Institute of Portland, Oregon, to make the call. In July 2006, the group laid down its verdict: the subspecies exists. The Fish and Wildlife Service is now digesting this ruling at typical government speed.

Although it might be tempting just to see a story of evil developers and good conservationists, the mice highlight a more fundamental problem. The act includes in its definition of species 'subspecies' and 'distinct population segments' but offers no definitions for either of these categories. The government agencies that enforce the act do provide guidance but the researchers called on by the government to determine objectively whether various groups of organisms are listable are mostly left to their own devices. Devices that have, over the years, littered the scientific literature with at least two dozen definitions of the concept of a species.

"We have more definitions than I can even remember," says Scott Steppan, a molecular systematist at Florida State University in Tallahassee, and a member of the deciding panel at the Sustainable Ecosystems Institute. Although most biologists agree that species are real entities — they exist without humans around to assign them — the distinctions are not clean-cut.

The reason why the relation between this mouse and that mouse can't be nailed down with jurisprudential exactitude is the same reason that there are any mice in the first place: evolution. The problem is that the idea of a distinct species predates Darwin's insights into their origins. Carl Linnaeus thought that species were made separate from one another by God, and that they stayed that way. But Darwin showed us otherwise. As one species splits into two over the millennia, there is no magic generation in which they are clearly separate. "It is kind of like asking when you are a child and when you are an adult — where is the boundary?" says Steppan.

Splitting hairs

Georgina Mace, head of the population biology programme at Imperial College, London, and two of her colleagues have chronicled how the number of species has changed under the rubric of 'taxonomic inflation'[4]. "In well studied groups, the number of species is increasing very rapidly, and that is in large part from the elevation of subspecies to the species level," says Mace.

According to their analysis, subspecies are becoming species not just at an increasing rate, but also and rather more problematically at varying rates in different parts of the natural world. For example, ant taxonomists have decided that anything that's worth separating should be separated at the species level, and have no truck with subspecies at all. Butterfly taxonomists, however, like the triple-barrelled name approach and dote on subspecies. As a result, the numbers of ant species and butterfly species are not directly comparable.

The implications of taxonomic inflation for conservation are wide-ranging, from increasing the number of endemic species in well studied areas and heating up 'hotspots' to making it almost impossible to figure out whether rates of extinction are slowing down or speeding up. Severe inflation can also, as in economics, lead to devaluation: if the smallest distinctions are raised up to the level that defines a species, the idea of a species loses some of its power.

S. & D. & K. MASLOWSKI/FLPA

The meadow jumping mouse (*Zapus hudsonius*) is rampant throughout the United States.

What is driving the inflation? According to Mace, much of it can be traced to the adoption of the phylogenetic species concept over the older biological species concept. In the biological species concept, if two fertile creatures cannot produce fertile offspring then they're not the same species, although there is some wiggle room for groups that have rare hybrids but keep their gene pools more or less separate.

The phylogenetic species concept rests on the idea of diagnosable differences. If one population has a particular characteristic: a red head, say, or a particular curve of a bill — and the characteristic can be shown to be inherited, it is taken as evidence of a unique evolutionary history, which would qualify the population as a distinct species. By the phylogenetic species concept, any difference that can be placed on a common limb of a family tree could count as a separate species, but how far up the branch the pruning shears are applied does, in practice, vary between taxonomists. "We can all agree on the data, but we can't all agree on how to apply the names," says Jack Dumbacher, a molecular ecologist at the California Academy of Sciences in San Francisco.

Lack of definition

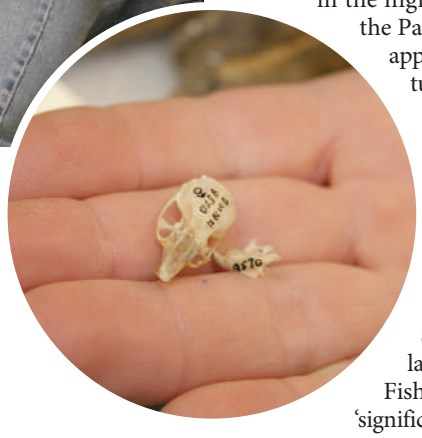
The phylogenetic approach can lead to a preference for splitting over lumping. It can also call older taxonomy into doubt, as in the case of the polar and brown bears. Although the two types of bear are distinct populations that lead different lives, studies of mitochondrial DNA suggest that brown bears do not share a common ancestor that does not also have polar bears as descendants; if you want to find one point on the tree from which all brown bears branch out, you will have to accept polar-bear branches in the same cluster.

By at least one reckoning, the phylogenetic approach comes up with 48% more species than the biological species concept does for the same group of organisms[5]. More species mean smaller groups and smaller ranges, so the groups are more likely to qualify as endangered. If subspecies are ascending to the species level, it may be reasonable to assume that heretofore un-named sub-subspecies — the sorts of thing that the Endangered Species Act already recognizes as 'distinct population segments' — are rising to the level of subspecies, and so on up the chain. And that may secure their conservation more attention.

"We have more definitions for a species than I can even remember." — Scott Steppan

The fact that it is now so easy to get gene sequences has contributed to the new ascendancy of phylogenetic over biological concepts. "We are able to slice the genetic pie thinner and thinner," says Craig Manson, who was assistant secretary for the Fish and Wildlife Service and the National Park Service from 2002 to 2005. Now teaching law at the University of the Pacific, McGeorge School of Law, Manson sees the Endangered Species Act as a creation of its time — a time when genetic data were still scarce and no one could fore-

J. GLADER



Rob Ramey argues that Preble's mice are no different from the common meadow jumping mouse.

see the coming orgy of re- and sub-classification. In today's world, he says "the act is sort of working — [but] not very well". The act has no clear thresholds below which a group of organisms is not considered a unit for protection purposes. "I think there needs to be a conference at the national level with the best experts in the scientific community that can be found, and let's hear this issue," says Manson. "It was one of the things I intended to do, and I just didn't get it done."

However, one corner of the United States has gone quiet over the definition for the tricky category of 'distinct population segment' — the smallest taxonomic unit that is listable. In 1991, Robin Waples of the National Marine Fisheries Service in Portland, Oregon, came out with his idea of how the term ought to be defined, and it stuck, at least for his agency[6]. 'National Marine Fish' shares the job of enforcing the Endangered Species Act with the Fish and Wildlife Service, and deals with the anadromous creatures — those that divide their lifecycle between salt and fresh water — as well as the truly marine animals. Although this might seem a smallish niche, in terms of the Endangered Species Act it's a big thing; of the ten listed populations in the United States on which most money is spent, eight are anadromous salmon or steelhead (rainbow trout). Anadromous fish nearly always return to their natal stream to reproduce, and so fish from different streams, even though they might brush scales in the wide ocean during their rambling years, are almost completely isolated when it comes to breeding. Waples requirements for a distinct population segment are: "it must be substantially reproductively isolated from other conspecific population units", and "it must represent an important component in the evolutionary legacy of the species".

So in the case of the chinook salmon (*Oncorhynchus tshawytscha,* which, as it happens, is considered by many to be the tastiest), those that spawn in the Columbia River basin are divided up into at least eight distinct population segments on the basis of their specific location and the timing of their runs. Some of the salmon travel to the ocean in spring, others in the summer and autumn. Although there are many more than eight populations in the basin, their listing groups them together into units of a manageable size.

Micromanagement

As Waples describes it, the distinct population segment combines a biological description of the relationships with a clear value judgement as to whether a population is important. "You really need something besides science to decide where on that level you are going to focus," he says. "There are hundreds of thousands of distinct stocks. It was not the intention of the framers of the Endangered Species Act to micromanage at that level."

Surprisingly, the gambit of defining a unit with a level of value judgement did not result in a rush of litigation, even in the highly fraught conservationist world of the Pacific Northwest. Waples says that the approach has become enough of an institution that it would take "a pretty compelling case" for things to change. As a result, although regulated industries and environmentalists might square off on other matters, such as how listed fish are going to be protected, the distinct population segment does not come up.

But although the Fish and Wildlife Service signed on to similar guidance language to that used by National Marine Fish — its guidelines on 'discreteness' and 'significance' are modelled on Waples' approach — experts mostly agree that it is less likely to list groups with only shallow genetic distinctness. "On this issue," says Manson, "they have never been in agreement, and they have grown further apart."

Without agreed ways of making judgement calls, personal feelings about conservation in general and about individual organisms in particular, can end up influencing ideas about classification. Ramey, the man whose genetic analysis nearly de-Prebled Preble's meadow jumping mouse, says that his differences with his detractors are "conceptual and philosophical". Conceptually, he says the fact that when at the museum (he now works independently) he had defined a threshold for what he would consider a subspecies before he did the analysis is proof that he was taking an unbiased look at the mouse. The others, he says, disagreed with his thresholds because they started off wanting the mouse to be a subspecies. But the report by the Sustainable Ecosystems Institute concentrates

not on concepts but on data, and rejects much of Ramey's evidence as being based on insufficient and poor quality data and contaminated samples. Ramey replies that his evidence does not hinge on those data points alone. After a slightly sharp series of exchanges on the Ramey work in the pages of *Animal Conservation*, experts seem to remain divided over the case. "If we look at the history of taxonomy, we often see that there have been many species named and then later synonymized," adds Dumbacher, who also sat on the Preble's mouse panel. "It may be that in the future, when the right studies are done, Preble's jumping mouse might be synonymized."

Criticisms of his data notwithstanding, Ramey says that his threshold is the right one, and that the level of difference between the Front Range mice and the rest is just too slight, in a world of limited resources, to justify any strenuous efforts to protect them. "We have to be willing to set priorities and stick to them. Right now, if you look at it, everything is a priority," he says.

Some suggest that the answer to all these problems is to ditch the taxonomic approach and shift to a totally different model of conservation law, such as ecosystem-based conservation. Many conservationists are thinking about concepts such as putting value on ecosystem services, such as water filtration and carbon sequestration. Mace sketches one possible ecosystem conservation model, in which several species' ranges are overlaid until a clear unit of space emerges. "You may have to deal with extreme specialists and those who like the edges of ecosystems separately," she adds. But when asked whether she thinks an approach like this will catch on in the near future, she sighs. "No," she says. "I think it is going to be very difficult because of the amount of expectation in policy-makers and legislators about the reality of the species concept. They really believe in it." ■

"We can all agree on the data, but we can't all agree on how to apply the names." — Jack Dumbacher

Emma Marris is a correspondent for *Nature* in Washington DC.

1. Barnes, P., Matheus, P., Shapiro, B., Jensen, D. & Cooper, A. *Science* **295,** 2267–2270 (2002).
2. Ramey, R. R., Liu, H.-P., Epps, C. W., Carpenter, L. M. & Wehausen, J. D. *Anim. Conserv.* **8,** 329–346 (2005).
3. Arbogast, B. S., Dumbacher, J. P. & Steppan, S. J. Evaluation of Scientific Information Regarding Preble's Meadow-Jumping Mouse (Sustainable Ecosystems Institute, Portland, Oregon, 2006); available at http://mountain-prairie.fws.gov/preble/Prebles_SEI_report.pdf.
4. Isaac, N. *et al. Trends Ecol. Evol.* **19,** 464–469 (2004).
5. Agapow, P. *et al. Q. Rev. Biol.* **79,** 161–179 (2004).
6. Waples, R. S. *Mar. Fish. Rev.* **53,** 11–22 (1991).

The big name hunters

Professional taxonomists often bristle at non-professionals who name new species without going through peer review. But are amateur naturalists really bad for science? **Brendan Borrell** reports.

The death adders of Australia are not adders at all. Their closest relatives are cobras and coral snakes, but early naturalists were fooled by the snakes' stout body and triangular head. Even today, their taxonomy is a riddle: no one really knows where one species of death adder ends and the next begins.

In the late 1990s, only three species of death adder had been recognized, but herpetologists suspected that there were at least twice as many. Ken Aplin, then a curator at the Western Australian Museum in Perth, had spent years collecting data to back up that hunch. But before his study could be published, Raymond Hoser, a herpetologist not affiliated with an academic institution, described five new species of the snake in a 1998 issue of *Monitor*, a hobbyist magazine he edited for the Victorian Herpetological Society. Under the taxonomic code of the International Commission on Zoological Nomenclature (ICZN), Hoser's names — printed and disseminated to society members — take priority over any subsequent descriptions of the species. Aplin had been scooped[1].

In the competitive world of taxonomy, countless amateurs have found success by collaborating with academics. Many professionals welcome their contributions; amateur enthusiasts are, in essence, a free workforce at a time when funding for basic taxonomy is waning. But cases such as Hoser's make some scientists wary of such contributions.

Hoser, who runs the snake-removal service Snakebusters in Melbourne, paints the picture as a classic case of academic élitism. "The description of me as an amateur is complete rubbish," he says. "There's no one in history who has spent so much time dealing with, looking at, catching and breeding death adders as myself." But his critics say it is not Hoser's credentials that they challenge. "A steady drip of shoddy descriptions" is how Wolfgang Wüster, an evolutionary biologist at the University of Wales in Bangor, describes Hoser's work. In a published critique, he and other leading herpetologists argue that Hoser "almost invariably fails to provide adequate information on his species, on their types, or on the material he has examined", making it difficult to repeat and test the observations[2]. Hoser, for his part, says that his descriptions contain more than adequate information.

Seek, locate, describe

But the ICZN, the group in charge of setting the ground rules for taxonomy, says it cannot police the quality of every published description. "It's a very tricky area to work in," says Andrew Polaszek, executive secretary for the organization. The commission, he says, will arbitrate only on pure nomenclature issues. In such a dispute, it will assess whether a Latin name put forward for a new species is valid under the taxonomic code. The code states that authors need to print the description of their species on paper, designate a type specimen, and list features that distinguish it from others. So just because a description is valid doesn't mean it is good. "The commission does not like to get involved in subjective taxonomy," Polaszek says.

Some taxonomists have proposed that the ICZN change its rules so that new species can be described only in peer-reviewed journals or through some other formal

accreditation process. But others think that might impose too much of a burden on small society journals where many good descriptions are published without peer review. In short, no clear solution is forthcoming.

Centuries ago, amateur was the only way to be a naturalist. The linnaean nomenclature system arrived in England in 1760 and soon became "a parlour game for the leisure class", says retired historian David Allen in Winchester, UK. "People armed with guides to the flora took them out on their fashionable wanderings around the countryside." The surge in popular interest fuelled a rash of publications in club newsletters and self-financed monographs describing species and rearranging classifications. The backlash began almost immediately; even Charles Darwin criticized the vain "species-mongers" who perpetuate a "*vast* amount of bad work"[3]. He thought the problem stemmed from this notion of priority, which he called "the greatest curse to natural history".

R. HOSER

Snakes alive: Raymond Hoser has devoted years to studying Australian death adders.

As a case in point, consider orchids. Two specialists recently published a plea against what they call "taxonomic exaggeration"[4]. As European botanists continue to subdivide orchid species that are not genetically distinct, they artificially inflate both the diversity and rarity of local flora, which may shift conservation priorities away from remote areas with fewer described species.

A question of honour

Orchids have long attracted a plethora of amateurs. In 2002, for instance, Michael Kovach smuggled a ladyslipper orchid from Peru and asked that a taxonomist at the Marie Selby Botanical Gardens in Sarasota, Florida, name it after him. Selby's experts reportedly knew that Eric Christenson, an unaffiliated taxonomist also in Florida, had his own description of the species scheduled for a forthcoming issue of *Orchids*. Selby rushed a two-page description of *Phragmipedium kovachii* to print as a supplement to its house journal. Kovach eventually pleaded guilty to illegal possession and trade of an endangered species; Selby was fined for its role in the scandal. None of this matters in the eyes of the taxonomic code, which will honour Kovach for ever.

Such 'orchid fever' has coloured the work of more than a few budding taxonomists. "There are lots of orchid amateurs who do not have a taxonomic background and who once in a while feel like they can make taxonomic changes," says Calaway Dodson, emeritus curator at the Missouri Botanical Garden. "Generally, it's a disaster." But some amateurs, he says, do excellent work. In Florida in the 1960s, a retired surgeon called Carlyle Leur began photographing the orchids of the state, and went on to document the entire country. Leur has since published more than 30 painstakingly illustrated monographs on the Pleurothallidinae, a daunting subfamily of orchids that contains more than 3,500 species.

Perhaps, then, the energy of amateurs can best be harnessed to fill the gaps in areas where few professionals work. The number of both professional and amateur taxonomists has been declining in Britain since the 1950s, according to a survey done in 2002 (ref. 5). But because they are distributed so widely, amateurs can generate better geographical coverage of flora and fauna, and focus on more descriptive taxonomy, leaving professionals free for molecular systematic studies.

Ground force

One successful venture, launched in 2002, is a partnership between government agency Natural England and the Natural History Museum in London, to encourage amateurs to contribute their data to local recorders. Museum experts have trained fly-fishermen to identify river flies and have enlisted members of the Ramblers' Association to monitor mature elm trees during their walks. The project met resistance from both sides at first, says sociologist Claire Waterton of Lancaster University. Some amateurs were reluctant to submit their data because they didn't understand how the work would be used, whereas others were worried that they would be judged unfavourably by professionals if they made mistakes. In turn, many professionals expressed concern about the quality of the data they might get. But in February, the British Mycological Society uploaded a database of fungal records to the

National Biodiversity Network. Compiled by local groups in Britain and Ireland, the database contains 140,000 recorded samples dating back to the eighteenth century.

Technology is also an important way to harness amateurs' contributions. Charles Godfray, an evolutionary biologist at the University of Oxford, UK, has worked to set up a peer-reviewed, single repository for all taxonomic information online — a sort of wiki-taxonomy[6]. That, he says, could help amateurs check the taxonomic designations of species that have been described. "The single thing that stops amateurs from being better involved in the process of taxonomy is getting at the literature," says Godfray. He is beginning to see his dream realized with a test website called CATE, for 'creating a taxonomic e-science', which he hopes will hold the taxonomy for aroids (popular house plants) and hawkmoths.

Polaszek notes that there is plenty of work for both amateurs and professionals. "We've got tens of millions of species to be described, and the easier this is, the better it is for everybody," he says. The ICZN is setting up a new system, called ZooBank, that requires species descriptions to be registered online. Within a year, he says, ZooBank could even be modified to include purely web publications such as CATE.

"We've got tens of millions of species to be described, and the easier this is, the better it is for everybody." — Andrew Polaszek

But could such changes sort out the mess over Australian reptiles? Perhaps not. In the 1980s, two amateur herpetologists called Richard Wells and Ross Wellington published more than 550 species descriptions that have since been changed[7]. An attempt to annul the work of the pair was rebuffed by the ICZN and taxonomists still have to sort through this work to determine whether the names chosen by Wells and Wellington have priority over other publications.

Hoser, for his part, found inspiration in their example and christened one death adder *Acanthophis wellsi*. It may have been a fitting tribute, as the name itself was improperly constructed. In a redescription of the species, Aplin amended the name to *Acanthophis wellsi*. ■

Brendan Borrell is a freelance science writer in Arizona.

1. Aplin, K. P. *Monitor* **10,** 104–109 (1999).
2. Wüster, W., Bush, B., Keogh, J. S., O'Shea, M. & Shine, R. *Litteratura Serpentium* **21,** 67–91 (2001).
3. Burkhardt, F. H. & Smith, S. (eds) *The Correspondence of Charles Darwin, Volume 4: 1847–1850*, 210 (Cambridge Univ. Press, Cambridge, 1988).
4. Pillon, Y. & Chase, M. W. *Conserv. Biol.* **21,** 263–265 (2007).
5. Hopkins, G. W. & Freckleton, R. P. *Anim. Conserv.* **5,** 245–249 (2002).
6. Godfray, H. C. J. *Nature* **417,** 17–19 (2002).
7. Thulborn, T. *Nature* **321,** 13–14 (1986).

The royal raccoon from Swedesboro

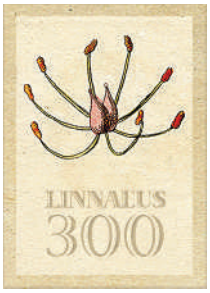
Although Linnaeus is best known for his botany and taxonomy, he was also an anatomist — and a keeper of pets. **Henry Nicholls** tells the story of Sjupp the raccoon.

Not many people would respond to the death of a pet by dissecting it. But Carl Linnaeus was an exceptional man. In 1747, Sjupp, his pet raccoon, clambered over a fence at the botanical garden in Uppsala and met a dog on the other side. The meeting was not a happy one for the raccoon. Unwilling to forego the opportunity to describe the raccoon's anatomy and find out where it sat in his system of nature, Linnaeus laid the mauled body out on a slab and picked up his scalpel.

Linnaeus's subsequent account of his raccoon is a perfect illustration of his powers of observation and attention to detail. But it is also a record of tenderness and affection, steeped in the rhythms of daily life at the botanical gardens. It deals with the dead animal's character as well as its anatomy — and in so doing reveals something of the character of the anatomist himself.

Sjupp was a gift from crown prince Adolf Fredrik, known to Swedish schoolchildren as 'the king who ate himself to death' because he keeled over in 1771 after putting away 14 helpings of a traditional pudding. He was also, admittedly less memorably, a keen amateur naturalist as was his wife. The pair acquired thousands of natural-history specimens, according to Anthea Gentry, a research associate at the Natural History Museum in London who has been studying the mammals and birds in the Swedish royal collections.

"For the royal couple, this was an enterprise driven by a desire to establish a collection of the most rare, conspicuous and interesting species," says Gentry. The queen invited



Linnaeus to catalogue and describe what they had acquired. Because they contain the specimens that Linnaeus used to describe various species — the 'type specimens' — the collections took on tremendous scientific significance. "They contain so many type specimens that they are as important as any other collection of similar size," says Gentry.

Sweet tooth

After Sjupp made the journey from the small zoo in the Royal Gardens in Stockholm to Uppsala he was kept in the royal manner to which he was accustomed. Although he would eat just about anything, "what he liked best were eggs, almonds, raisins, sugared cakes, sugar and fruit of every kind", Linnaeus observed[1]. "Should there be any cake or sugar on the table or in a cupboard he was on it in a flash, and thoroughly enjoyed himself. If a student came in who happened to have raisins or almonds on him, he at once attacked his pocket and fought until he had captured the spoil. On the other hand, he couldn't bear anything with vinegar on it, or sauerkraut, or raw or boiled fish."

Linnaeus went on from Sjupp's tastes to his temperament. "He became very friendly with people when he got to know them, letting them pat and play with him (especially if they ingratiated themselves by means of a few raisins)." But the raccoon also had a moody side. "Anyone who had once quarrelled with him found it almost impossible to get back again into his good books." Linnaeus's head gardener, who had once panicked and flapped when Sjupp bounded up to him and began to search his body for a tid-bit, suffered this disdain.

M. JONSSON

Pet portrait: Linnaeus hung this watercolour of Sjupp the raccoon at his summerhouse.

"From that moment Sjupp developed an irreconcilable hatred of the man," he wrote. "Every time he smelt him he began making a noise like a seagull, the sign that he was extremely angry." In short, Linnaeus concluded, Sjupp could be "as obstinate as a knife grinder".

His rich account of Sjupp's life reveals the many levels at which Linnaeus engaged with the natural world, says Karen Reeds, historian of science and guest curator of a tercentennary exhibition on Linnaeus and America at the American Swedish Historical Museum in Philadelphia. "Nature really captivated him emotionally as well as scientifically," she says. Testimony to his lasting affection, says Marita Jonsson, author of a book[2] on the power of place in Linnaeus's work, is a watercolour of Sjupp that hung in the study at Hammarby, Linnaeus's summerhouse just outside Uppsala.

Change of status

This fondness, though, held no squeamishness: the meat of his 1747 paper is an organ-by-organ description of the raccoon's anatomy. He named the raccoon *Ursus cauda elongata*, 'the long-tailed bear'. But by the time he published his tenth edition of *Systema Naturae* in 1758, he had revised its name to *Ursus lotor*, 'the washing bear', in the light of new evidence and to bring it in line with his binomial nomenclature system. "It was typical of the way he was constantly tweaking his classification scheme," says Reeds. "One of the most impressive things about Linnaeus was his readiness to change his mind as new information reached him."

Today's taxonomists are still struggling to come to an agreement about raccoons, mainly because of an overenthusiastic bout of 'splitting' in the early twentieth century. In a matter of decades, scientists had described more than 20 subspecies of the common Northern raccoon (Linnaeus's *Ursus lotor*, and now called *Procyon lotor*). Things got particularly out of hand in the Caribbean, where each island population was designated as a distinct species.

BRIDGEMAN ART LIBRARY

Linnaeus's study at Hammarby, Uppsala, Sweden.

A combination of molecular and conventional morphological approaches has now revealed that the Caribbean raccoons are probably recent introductions to the islands from the mainland population, and therefore undeserving of any special taxonomic status, allowing taxonomists to collapse the apparent diversity of raccoons back down to something more meaningful[3].

"In the Bahamas, they were delighted. They instantly changed the raccoon's status from endangered to invasive species and set up a control programme to eradicate them," says Don Wilson, a taxonomist at the National Museum of Natural History in Washington DC. But in Guadeloupe, the locals were far from happy, having taken some pride in the distinctiveness of their raccoon. "They love the little rascals," says Wilson.

Sjupp himself was no Caribbean exotic; he almost certainly came from 'New Sweden', the Swedish colony on the Delaware River founded in the seventeenth century. Pehr Kalm, one of the 'apostles' Linnaeus sent out into the world to gather its riches, went there in 1748 and reported that raccoons, or more precisely their skins, were an important part of the North American economy. "The hatters purchase their skins, and make hats out of their hair, which are next in goodness to beavers," Kalm noted in his *Travels into North America*[4]. "The tail is worn around the neck in winter and therefore is likewise valuable." The village Kalm spent his winters in was actually called Raccoon at the time; today, though, it is known as Swedesboro.

Reeds says that Sjupp almost certainly came from this area, as did many furs exported to Sweden and the Netherlands: "It seems plausible that they would have sent a live specimen along with them." When Kalm eventually returned to Uppsala, "Linnaeus was tremendously excited about what he found out", says Reeds. Linnaeus was roused from his sickbed where he'd been suffering from a severe attack of gout and returned to his *Species Plantarum* with renewed enthusiasm, she says. And at some point, he got his hands on a second live raccoon — there is one listed in the 1769 inventory of Linnaeus's menagerie. Whether it squawked at gardeners and snuffled through students' pockets, though, we do not know. ■

Henry Nicholls is a freelance writer in London.

1. Blunt, W. T. *The Compleat Naturalist: A life of Linnaeus* (Collins, London, 1971).
2. Jonsson, M. *Carl von Linné, Boningar, trädgårdar och miljöer.* (Bokforlaget Forum AB, 2003).
3. Helgen, K. M. & Wilson, D. E. *J. Zool.* **259,** 69–76 (2003).
4. Kalm, P. *Travels into North America* (Trans. Forster, J. R.) (Lowndes, London, 1772).

Publications should include an animal-welfare section

SIR — A large majority of the public supports the principle of animal experimentation to improve biological knowledge, human and veterinary health, nature conservation and, last but not least, animal welfare. This support, however, depends on strict adherence to the 3Rs (replace, refine, reduce) principle, described in a 2005 report by the Nuffield Council on Bioethics (see *Nature* **435,** 392; 2005). This is aimed to minimize animal numbers, pain, suffering and lasting harm.

A recent set of News Features on animal research ("A matter of life and death" *Nature* **444,** 807–816; 2006) identified considerable scope for advancing the 3Rs, which crucially depends on an effective spread of relevant knowledge and techniques. Several specialist journals — such as *Laboratory Animals*, *Lab Animal* and *ALTEX* — publish research aimed at advancing the 3Rs criteria, but their readership and impact are limited. Regulators and animal-care committees do not have the means to broadcast and implement novel techniques effectively. Many useful refinements are developed and applied by single labs, but do not get published because they were not a primary focus of the research or because they are not sufficient for a paper on their own. Much relevant information might thus never become widely available within and between these communities.

Most scientific journals require a statement of adherence to legal and institutional animal welfare guidelines. Some have their own codes of practice and ethical committees to guarantee high animal-welfare standards in published material (see, for example, *Animal Behaviour* at http://asab.nottingham.ac.uk/ethics/guidelines.php).

Journals could play a much more effective part, however, by including a 3Rs section in the methods section of published papers. First, this would allow authors of controversial papers to detail their measures to minimize pain, suffering and lasting harm. Second, it would let them describe novel tools or techniques used in the paper that serve the 3Rs. Journals could make this 3Rs section optional, and — depending on the significance and length of the material — could either include it in the print version of the paper or make it available online only as accompanying supplementary material.

Leading journals such as *Nature* could pioneer such a policy. This would be in the best interests of editors, scientists and the public, as well as to the benefit of the experimental animals.

Hanno Würbel
Division of Animal Welfare and Ethology, Institute of Veterinary Physiology, Justus Liebig University of Giessen, 35392 Giessen, Germany

***Nature*'s editorial policy on papers reporting animal experiments is at www.nature.com/authors/editorial_policies/experimental.html. We welcome readers' comments on this proposal at Nautilus, our author blog: http://blogs.nature.com/nautilus/2007/03/proposal_for_journals_to_inclu — Editor, *Nature*.**

Scientists need to confront economists about peak oil

SIR — Your News Feature "That's oil, folks" (*Nature* **445,** 14–17; 2007) highlights the debate over depletion of the world's oil reserves. I would like to make some additional points.

First, the proponents of the peak-oil theory are predominantly *Nature*'s constituency — scientists — whereas the vocal opposition are, to a significant extent, economists. They seem to believe that the geological reality of finite conventional oil resources and the thermodynamic constraints on energy production from alternative hydrocarbon sources can be overcome by a sufficiently high price signal.

Second, there are many statistical and energy-production data supporting predictions of imminent energy decline. For example, a chart of annual discoveries of oil during the twentieth century shows that, despite tremendous advancement in discovery and extraction technology during this period, oil discoveries have been on a downwards trend for nearly 50 years (see *ASPO Newsletter* 73; January 2007). Although huge, non-conventional oil resources exist — for example: tar sands, shale oil and even biofuels — harvesting these resources is likely to produce little or no energy profit.

Third, scientists warning of energy decline are seriously disturbed by this issue, for many reasons. One is the annual increase in the world's human population. Until recently this has been sustained by increasing grain production, made possible by the oil-driven 'green revolution'. However, grain consumption now exceeds production and reserves are dwindling rapidly. The availability of food will be further eroded by the diversion of grain to production of biofuel.

Most people lack sufficient scientific training to appreciate the strong evidence for, and dire consequences of, an imminent decline in oil production. They are easily lulled into complacency by those with a vested interest in delaying any mitigating responses. The scientific community must unite behind the issue of energy decline.

Michael Lardelli
School of Molecular and Biomedical Science, University of Adelaide, North Terrace, Adelaide, South Australia 5005, Australia

Concept of a bacterium still valid in prokaryote debate

SIR — Nigel Goldenfeld and Carl Woese, in their Connections Essay "Biology's next revolution" (*Nature* **445,** 369; 2007), seek a change in concepts of 'organism, species and evolution' because of the prevalence of lateral gene transfer among bacteria. However, it has been clear for half a century that biological species are epiphenomena of sex, and exist only in sexual eukaryotes — but not in bacteria, which transfer genes laterally without sexual cell fusion. The essay exemplifies a common linguistic confusion caused by those who wish to equate microbes and prokaryotes (see W. Martin and E. V. Koonin *Nature* **445,** 21; 2007).

This state of affairs acutely highlights the continued need for the classical 'concept of a bacterium' (prokaryote) put forward by R. Y. Stanier and C. B. van Niel (*Arch. Mikrobiol.* **42,** 17–35; 1962). The inapplicability of biological species claimed for 'microbes' generally is emphatically not true for sexual protists (eukaryotic microbes). There are even more positive characters shared by all prokaryotes (for example, chromosomes attached to surface membranes) than noted by Martin and Koonin. The only circumstance in which we could reasonably abandon the term 'prokaryote' would be if the biological community as a whole were to accept the traditional use of 'bacteria' instead to embrace both archaebacteria and eubacteria.

The real need is not for a 'revolution' in language and change in the classical concept of an organism, but for molecular evolutionists to make a more serious attempt to understand it. Organisms are not mere assemblages of genes, whether inherited vertically or laterally, but cells (or integrated assemblies of cells) in which there is a mutualistic cooperation of genomes, membranes, skeletons and catalysts that together make a physically and functionally coherent unit capable of reproduction and evolution.

Thomas Cavalier-Smith
Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

Faulty logic

SIR — Gregory C. Beroza, in his Book Review of *Richter's Scale* by Susan Hough (*Nature* **445,** 599; 2007), observes that Hough is on "shaky ground" when she uses extrapolation to explain Charles Richter's "many idiosyncrasies and indiscretions".

Seismologists, including Hough, are surely familiar with being on shaky ground.

Alex C. W. May
Northwest Institute for Bio-Health Informatics, University of Manchester, Oxford Road, Manchester M13 9PT, UK

COMMENTARY

Linnaeus in the information age

As we celebrate the visionary genius of Carl Linnaeus, it is time to analyse how professional taxonomy interfaces with the rest of biology and beyond. Where next for Linnaeus's heirs, asks **H. C. J. Godfray**?

It is one of the triumphs of contemporary science that we have a means of naming and referring to all described organisms on Earth, as well as their fossil ancestors. It was Carl Linnaeus who realized that, to understand anything in science, things have to have a name that is recognized and is universal — a lesson that has not been forgotten by the creators of today's gene and protein databases. The year in which we celebrate Linnaeus's achievements is an opportune time to think about how modern taxonomy relates to the rest of biology, and how this relationship may evolve.

In Linnaeus's time, the natural world seemed static and relatively untouched by humans. This is no longer true, and a greater appreciation of our planet's changing biodiversity is increasing the demand for taxonomic information and expertise. Simultaneously, the nature of biology is changing, so that tasks that were once the unique preserve of the taxonomist can now be done by most biologists. In confronting these realities, taxonomists need to think about where to prioritize their efforts, and how to make the best use of limited resources.

BRIDGEMAN ART LIBRARY

Do it yourself

Most biologists use taxonomic information in one way or another, from the dedicated molecular biologist who is content that 'the fly' is understood to mean *Drosophila melanogaster*, to the community ecologist who has to identify hundreds of species at a study site. Taxonomy provides opinions on species boundaries, and on the phylogenetic relationship between species. It provides a stable naming system that in today's jargon is a portal to a huge, if not always easy to access, store of information about a species. Advocates of increased investment in taxonomy highlight all these potential benefits, but often forget that many of the traditional user communities of taxonomy can now investigate these questions themselves.

Consider the role of taxonomy in deciding what is a species. Arguments for the importance of taxonomy frequently note how accurate species identification is essential for, say, combating a particular human pathogen, or for discovering the appropriate biological control agent for an agricultural pest. In the past, these critical decisions could be made only by taxonomists, and typically involved careful morphological study and comparison of preserved specimens. Today, molecular markers are increasingly used to study population differentiation, and most biologists can do this. Indeed, for any taxonomic issue with medical or economic consequences, it is inconceivable that molecular techniques would not be applied. Similarly, biologists who might have previously looked to taxonomists to provide a phylogeny of a group are finding it increasingly easy to do it themselves as sequencing becomes cheaper and more widely available.

Biodiversity scientists, those interested both in fundamental and applied questions, frequently need to identify large numbers of species. To do this, they need either the products of taxonomy — field guides, monographs and floras — or the taxonomists themselves. Excellent identification resources exist for some groups in some geographic regions, but often these are absent, especially in the tropical biodiversity hotspots that need most protection and research. Ten years ago the only solution to this would have been more investment in traditional taxonomy. Today there is an alternative: molecular identification methods, DNA barcodes or related technologies that would allow the high-throughput allocation of specimens to different sequence clusters. DNA barcodes could develop completely independently of linnaean taxonomy, although this is definitely not the intention of its leading proponents.


P. SCHERMEISTER/CORBIS

Ecologists in the field rely increasingly on molecular methods to 'do' taxonomy themselves.

So there is a threat that the usefulness of taxonomy may seem to diminish as new techniques become available. Taxonomy itself would remain intrinsically interesting, but the rise of alternative technologies could dangerously undermine arguments for increased investment in the subject. To counter this, I believe that taxonomists need to do the following: first, ask difficult questions about the efficiency and value for money of taxonomy as it is currently conducted; second, consider the needs of their user communities and prioritize some 'big science' projects to meet them; third, integrate their current practices with new molecular techniques, as many already do; and fourth, improve the ease with which taxonomic information can be accessed by non-taxonomists, and emphasize how taxonomy provides a portal to the totality of information available about different species.

More than a name

Some readers will dislike the utilitarianism of these arguments. Taxonomy is rightfully a proud science and does not need to justify its existence as a service provider. But it lives in the real science-funding world, and frequently suffers from it. Too often, taxonomists are pressurized by their bosses to reinvent themselves as ecologists, evolutionists or even computer scientists to chase research funding. Biology, and especially biodiversity science, need more taxonomists doing more real taxonomy. This will be easier to achieve if the field does today what Linnaeus did: provide its end users with products that are so good they clamour for more.

Linnaeus's brilliance was that he initiated the solution to the first bioinformatics crisis: how to organize knowledge about the increasingly large number of species that were being discovered during the age of exploration. His solution was appropriate to the days of paper publication and surface post, but what would his approach have been today? I believe he would have been a 'techie', exploiting the Internet and other modern means of sharing and coordinating data. No one denies that the web will be increasingly crucial for taxonomy, but what is much less clear is the degree to which taxonomy should transfer wholly to the web, and how this will change the hallowed rules of nomenclature enshrined in the International Codes. I think that a move to a universally accessible, web-only taxonomy is essential for the future health of the subject. But I also understand the caution felt by critics of these ideas.

Whether taxonomy does or does not move *in toto* to the web, I believe it is critical for the subject to emphasize its role as a curator of information about taxa, in addition to names and specimens. There are some excellent projects starting to do this, for example *FishBase*, which collates all types of information on its eponymous taxon. Here you can learn about a fish species' morphology, reproductive biology and ecology; the environments it can tolerate and its population dynamics; whether you can eat it and whether it can eat you. More projects like this are needed, along with moves to integrate basic taxonomy with new initiatives such as DNA barcoding.

"Linnaeus would have been a 'techie', exploiting the Internet and other modern means of coordinating data."

The web also has a role in making taxonomy more economically efficient. It can be used to collate the literature and specimen information required to do taxonomy, as well as disseminate its outputs. But why stop there? The type specimen plays a critical part in taxonomy, for example settling disputes over species identification. But type specimens are often very old and have deteriorated in some way. The more they are consulted the greater the risks of further damage, and access to types is sensibly restricted. Most types also lack associated sequence data. Perhaps we need a way to create a new form of type specimen (call it a cybertype) to be displayed on the web using the very best current imaging methods — often far superior to normal examination — and for which appropriate sequence data are available. Plant taxonomists can already designate a new specimen (an epitype) to help interpret an ambiguous original type, and I think it would be helpful to extend this practice to animals and specifically adapt the concept for web and molecular taxonomy.

Access for non-professionals

As scientists, we are rightly encouraged to engage more with the public. Taxonomy is one of the few subjects (astronomy is another) where non-professionals can make genuinely important contributions, and the shear scale of Earth's biodiversity makes this engagement essential. Difficulties in accessing taxonomic information act as a brake on research and reduce its quality. Moving as much taxonomy to the web as possible will enfranchize a huge amateur community and greatly aid professionals in developing countries, where most biodiversity hotspots are.

By its nature, taxonomy is a broad subject that finds difficulty in prioritizing research and investment. Big taxonomic projects have sometimes been too big and all-embracing — for example, proposals to catalogue all the species on Earth — and have failed to progress. Other projects are exceptionally valuable for taxonomists (for example, lists of names and synonyms) but are of less relevance to taxonomy's user communities. The field could usefully work with the broader biological community to define and prioritize a series of big taxonomic projects that are achievable over the next decade or so, much like the physics and astronomy communities do.

My personal wish list, biased towards terrestrial eukaryotes and living rather than fossil organisms, would include a web-based global flora, a complete phylogeny of the living vertebrates, and a comprehensive web-based taxonomic and identification resource (morphology plus DNA barcodes) for the world's macrolepidoptera. I chose macrolepidoptera (butterflies and large moths) because I think that this is the only large, ubiquitous invertebrate group that is relatively easy for non-specialists to identify, and so would be invaluable as a biodiversity-monitoring tool. Others would pick a different set, and it will undoubtedly be hard for taxonomists and other biologists to reach a consensus. But such a list, backed by the community and firmly linked to end-user demand, would be a wonderful lever for bringing new resources into the field. ■

H. C. J. Godfray is in the Department of Zoology, University of Oxford, South Park Road, Oxford OX1 3PS, UK.

See also pages 231, 247 and 254.

Spreading the word

Keeping track of new species names is a growing challenge for modern taxonomists. **Sandra Knapp**, **Andrew Polaszek** and **Mark Watson** make the case for electronic publication of scientific names.

When Carl Linnaeus was born in 1707, plants and animals were not given 'scientific' Latin names. Instead they were referred to by 'phrase names' consisting of complete descriptions in Latin, sometimes running into several paragraphs. By the time Linnaeus died in 1778, taxonomy was an established discipline and people communicated information about organisms using his simplified binomial system. These two-word 'trivial names' are what we now call the genus and species. Twenty years after Linnaeus's death, rumblings about the lack of stability in the system of plant and animal naming led to the Codes of Nomenclature that today govern the naming of all living things. Biology owes a great deal to the willingness of the taxonomic community to adhere voluntarily to these codes: names, and their consistent application, make biology a repeatable science.

Modern taxonomy has engendered much comment in the past few years[1]. Its products, descriptions and hypotheses concerning the identity and relationships of life on Earth, are rightly seen as increasingly important in documenting biodiversity. But the business of communicating the findings has outpaced the traditional means of publication. In our view, the products of taxonomy are too useful to the rest of biology to be relegated to a handful of obscure, privately published books or journals of limited distribution, both physically and intellectually. Taxonomy must not be left behind in the electronic age.

Each of the four codes (for animals, plants, cultivated plants and bacteria) applies a set of conventions to naming, which are governed in slightly different ways[2]. But all the codes share two central pillars: validity of publication and the principle of priority — whereby the first validly published name for a given taxonomic concept is the correct one. The rules governing publication are arguably the most important of all the codes' regulations, as it is through reference to publications that hypotheses are supported or rejected.

In Linnaeus's day, publication involved printing presses and personal payment; many of the works published in the eighteenth century survive only as single copies in a few libraries. Linnaeus himself published and distributed the dissertations of all his students, some 400 printed books in limited print runs. How different from the world of publication today, with its instantaneous electronic access to all corners of the globe.

S. KNAPP/NHM

In Linnaeus's day, new names were published in printed books, few of which survive today.

Up to speed

The difficulty of tracking down taxonomic publications has rightly been identified as the greatest impediment to the work of taxonomists and its use by other biologists. Old, rare publications are difficult enough to find, but today there are some 25,000 new names for organisms proposed every year. These new names are published in thousands of journals, many of very limited distribution; keeping track is an almost impossible task. How can taxonomy join the rest of science and take advantage of the speed and reach of electronic publishing?

"The products of taxonomy are too useful to be relegated to a handful of obscure journals."

Online community initiatives, such as the International Plant Names Index (IPNI, www.ipni.org), which began life as *Index Kewensis* with a legacy from Charles Darwin[3], exist for the recording and electronic dissemination of new names published elsewhere. Similar compilations for bacterial names, and the recent ZooBank (www.zoobank.org; ref. 4) and MycoBank (www.mycobank.org) initiatives will contribute a great deal towards making names available to all. However, none of the current Codes of Nomenclature allows publication of new names by electronic media alone.

Will this change? Periodic tweaks are made to the codes, usually to maintain stability in the naming system. The communities governing the codes are conservative and wary of sweeping changes that may cause more problems than they solve. But as publishing goes electronic, the landscape is changing slowly. The zoological code now allows for publication on CD instead of print, while the botanical code is recommending publication in peer-reviewed print journals with electronic distribution instead of in print-only journals[5]. In our view, print-dependence for the naming of new organisms slows taxonomy down.

Eventually, pure web-based taxonomies, such as those being developed by the CATE project (www.cate-project.org/) and the Planetary Biodiversity Inventory projects (www.actionbioscience.org/biodiversity/page.html) will radically improve access to the products of taxonomy. However, electronic media alone will not speed up publication. Rapid review is also critical for rapid publication. Today most,

but not all, print journals in taxonomy are peer reviewed, and as publication migrates to the web it is important that standards go up and we don't create an unregulated free-for-all. Peer review is not now an explicit condition of publication for new names in most of the codes. We believe electronic taxonomic publishing must include rigorous peer review, alongside fast and wide dissemination of results.

A culture of rapid, high-quality peer review is essential to keep a field moving, and taxonomy needs to move faster to contribute effectively to advances in biodiversity science. An example of a successful journal that publishes simultaneously online and in print is *Zootaxa* (www.mapress.com/zootaxa/). It has proved a popular and effective means for publishing new animal names, in part due to its electronic distribution, but also due to its quick turnaround. A paper can take up to two years from acceptance to publication in some taxonomic journals, whereas in *Zootaxa*, turnaround varies from weeks to months. The Linnean Society of London and Blackwell Publishing are discussing a similar initiative for botanical names.

LINNEAN SOC. LOND.

S. KNAPP/NHM

A secure means of archiving is essential to protecting both old and new taxonomic information.

In a safe place

There are other issues on which taxonomists will need to work with the publishing industry to find solutions. Important objections to electronic publication are the problems of permanent archiving and accessibility. The current system, paper archiving, is itself no guarantee of permanence: the ancient 'copyright' Library of Alexandria was famously destroyed by fire on several occasions. In recognition of this, the codes require that copies of new names are deposited in at least five major public libraries. This remains necessary until stable and permanent electronic publications are attainable.

The worldwide web is archived regularly (www.archive.org), but with no guarantee that this will be permanent. However, it is inconceivable that investments in electronic databases such as those for genome sequences or astronomical data will be allowed to disappear — a stable archive is protection of this investment, so there is ample incentive to get it right. Archiving is just as essential to protecting the investment made over centuries in taxonomy, and projects such as the Missouri Botanical Garden's Botanicus (www.botanicus.org) and the Biodiversity Heritage Library (www.bhl.si.edu/) will go a long way towards doing this.

An alternative to archiving all original publications is an automated system for registering new names, whether voluntary or mandatory. Names for fungi can now be registered on Mycobank, and from mid-2007 a form for registering new names will be available via ZooBank. The flowering-plant community will, however, continue to use IPNI as its recording system. These systems are all voluntary, or in the case of IPNI, maintained by institutions. Moving to mandatory registration would imply that any new taxon name could not be considered valid until it had both been published and registered, a two-step process that would require a change to the codes.

Mandatory registration of taxonomic names has proved controversial in the past, largely for sociological reasons having to do with how it would work, rather than objections in principle: one system of centralized registration was rejected by the 1999 International Botanical Congress. However, with the advent of globally unique and persistent identifiers for taxon names or web documents (a concept borrowed from software applications) by groups such as the Taxonomic Databases Working Group www.tdwg.org) and the Global Biodiversity Information Facility (GBIF, www.gbif.org), the pieces necessary for making name registration work effectively across all biology are coming together.

Open access

GBIF is also dedicated to making taxonomic (biodiversity) information accessible as widely as possible. Restrictions to online access of journals and articles can work against this, and although free access to the entire published literature may not be possible, a registration system that guaranteed open access to important nomenclatural information (such as names and original descriptions) could satisfy the needs of both taxonomists and the publishing industry. What better longer-term role for GBIF than to assume eventual responsibility for registration and assignment of globally unique identifiers for all new taxonomic names?

Ultimately, issues surrounding access to new names transcend the medium of their publication. The increasing reliance by peer-reviewed journals on electronic editions and the decreasing purchasing and storage capacities of libraries worldwide means that print media face an uncertain future. Access to taxonomic information will increasingly come to rely on the Internet, whether by 'unitary'[1] or by distributed means. As already noted, the codes currently accept publication of new names only in library print or CD copies, preferably those linked to major name-indexing centres, but the main dissemination of such works can, and we think should, be electronic.

The time has come to think as Linnaeus himself once did when he and his students began to use his binomial 'nomina trivialia' as aids to memory. Linnaeus knew his work was important, and that it mattered. He used whatever new ideas and technologies he could to disseminate it and make it useful. It would be fitting if in 2007, his tercentenary, the taxonomic community made significant steps towards bringing nomenclature into the electronic age. ■

Sandra Knapp is in the Department of Botany, and Andrew Polaszek is at the International Commission on Zoological Nomenclature, The Natural History Museum, Cromwell Road, London SW7 5BD, UK; Mark Watson is at the Royal Botanic Garden Edinburgh, 20a Inverleith Row, Edinburgh, EH3 5LR, UK.

1. Godfray, H. C. J. *Nature* **417,** 17–19 (2002).
2. Knapp, S. *et al. Phil. Trans. R. Soc. Lond. B* **359,** 611–622 (2004).
3. Nic Lughadha, E. *Phil. Trans. R. Soc. Lond. B* **359,** 681–687 (2004).
4. Polaszek, A. *et al. Nature* **437,** 477 (2005).
5. Knapp, S. *et al. Taxon* **55,** 2–3 (2006).

BOOKS & ARTS

Biological programming

The digital nature of molecules such as DNA means they can be used in computers.

Genesis Machines: The New Science of Biocomputing
by Martyn Amos
Atlantic Books: 2006. 320 pp. £18.99

Christoph Adami

When visiting the Santa Fe Institute in autumn 1994, I was chatting with Chris Langton when Stuart Kauffman came running down the hall waving a manuscript and shouting unprintables. When we cornered him, he was slapping his forehead mumbling: "I could have done this, if I had only thought of it!" It turns out that 'this' was the first example of a DNA-based computation, carried out by Leonard Adleman of the University of Southern California, launching the field of DNA computation (*Science* **266,** 1021–1024; 1994). Martyn Amos's book *Genesis Machines* looks back on the 12 years since this event, and speculates about the future of the increasingly intertwined fields of biology and computer science.

The computational feat reported in Adleman's seminal article was innocuous enough: examine a graph of seven nodes, and determine whether a one-way path exists that connects all the nodes once and only once (an example of the hamiltonian path problem). But the importance of this work did not lie in the sophistication of the problem, but in the fact that it showed that strands of DNA, mixed together in a vial, could be controlled such that their biochemistry could be viewed as a computation. And this is perhaps the central message that Amos tries to convey in his book: all physical systems can be viewed as performing computations; it is down to the skill of the investigator to make them perform useful ones.

Since 1994, DNA computation has advanced considerably, and programmable computers have been developed using DNA molecules only and with no moving parts. But for the reader to appreciate the story of how vials of DNA molecules could be coaxed to do something nature never intended, Amos has to reach far beyond the basic biochemistry of DNA, to the origins of computer science and complexity theory. Indeed, DNA computation was not the brainchild of biologists, but computer scientists who became fascinated by the 'digital' nature of DNA and its role in information processing. The roots go back to luminaries such as John von Neumann and Alan Turing, and Amos goes to considerable lengths to expose these foundations.

A. PASIEKA/SPL

Programmable computers have been developed that use DNA.

Perhaps one of the weaknesses of this book is a result of the unavoidable interdisciplinary nature of its subject. For the lay reader for whom this book is written, the concepts of Turing universality, the complexity of 'P versus NP', and the biochemistry of DNA are likely to be difficult subjects, and Amos's introductions include biographical sketches of their main protagonists. As a result, more than 100 pages go by before Adleman's pioneering experiment can be described in detail, and even then it is interrupted by a narrative of the trials, tribulations and ultimately triumph of Kary Mullis and his discovery of the polymerase chain reaction.

Nevertheless, the central discoveries in this rapidly evolving field are covered, and the inevitable critics of this endeavour are not overlooked. Amos does a good job of injecting some sobriety into the narrative, lest the reader believes that home computers running on DNA are just around the corner. Indeed, he points out right from the beginning that DNA computation is unlikely to replace its silicon-based forerunner, but rather be confined to special problems. What these will be is not at all clear, because the initial excitement about DNA's ability to tackle problems in the most difficult category, such as factoring a number into its prime components, quickly waned when it became clear that thousands of litres of DNA would be necessary to solve problems that are now just beyond the reach of conventional computers. But Amos insists that DNA computation is interesting in its own right because of the lessons it teaches us about computation in biology. The last two chapters are therefore devoted to endeavours only indirectly related to the quest for a DNA-based computer, namely DNA self-assembly and synthetic biology.

As Amos explains, the stability of DNA and its digital nature can be used to create — or should we say program — structures on a nanometre scale, a feat that is exceedingly difficult to achieve using non-biological material. Inspired by the pioneering work of crystallographer Ned Seeman at New York University, Erik Winfree and Paul Rothemund of the California Institute of Technology set out to show that DNA can be programmed to fold into arbitrary structures whose intricacy can only be revealed by atomic force microscopy.

Synthetic biology, on the other hand, concerns the creation of biological systems from scratch or from components that had other functions. Behind it lies the concept that we do not need to understand every single detail of a complex system in order to be satisfied with what we know. After all, very few people understand down to the level of the electronics the workings of my laptop computer. But this does not worry me because I know that someone knows how to make it, so there is sufficient knowledge to create another one just

like it. The same can be said for biomolecular systems: we may never be able to comprehend every molecular interaction within a cell, but if we can build a system that behaves just like its natural counterpart, then perhaps we should be satisfied with that.

Amos's account of the different paths to synthetic biology covers most of the recent advances that have made the headlines. These range from pieces of DNA called 'biobricks' and the 'repressilator' to a way of coaxing yeast to produce the precursor for a malaria drug — a feat for which Jay Keasling, a biochemical engineer at the University of California, Berkeley, won *Discover* magazine's first 'scientist of the year' award.

But I have to take Amos to task for presenting the work of James Shapiro, a microbiologist at the University of Chicago, as if his investigations of the mechanisms of evolution somehow contradict darwinian evolution. Shapiro, who studies the many ways in which organisms can actively restructure their genetic material, believes that such processes should be more widely incorporated into evolutionary thinking. Although Darwin could not possibly have anticipated such processes, they follow the general darwinian paradigm and do not, in my opinion, necessitate a 'third way' somewhere between creationism and what Shapiro calls 'neo-darwinian orthodoxy'.

But apart from this quibble, this is an enjoyable book that could perhaps have profited from more illustrations to convey some of the computational and experimental methods. I recommend it to anyone interested in computation writ large who is not afraid to cross disciplinary boundaries that once seemed impassable. ■

Christoph Adami is at the Keck Graduate Institute of Applied Life Sciences, Claremont, California 91711, USA.

Mammals on the move

The Beginning of the Age of Mammals
by Kenneth D. Rose
Johns Hopkins University Press: 2006.
428 pp. $150

Zofia Kielan-Jaworowska

Just looking at *The Beginning of the Age of Mammals* by Kenneth Rose, with its Henri Rousseau-style jacket picture, catapults the reader right into enchanted Early Eocene life. The book is a scholarly treatment of an important period in the evolution of mammals. Mammals originated some 225 million years ago, and for about 160 million years they were small nocturnal creatures living in the shadow of the dinosaurs. Most of them would have been size of a shrew or rat, with only a few being as big as a fox. When the dinosaurs died out at the end of the Cretaceous, 65 million years ago, mammals changed lifestyle and started to increase their size, with many entering niches abandoned by the extinct giants. This was the age of mammals.

Kenneth Rose's book *The Beginning of the Age of Mammals* is devoted to this explosion of mammals, particularly their evolution during the first half of the Cenozoic period. Rose highlights in particular the development of placental mammals in the Palaeocene and Eocene (between 65 million and 33 million years ago).

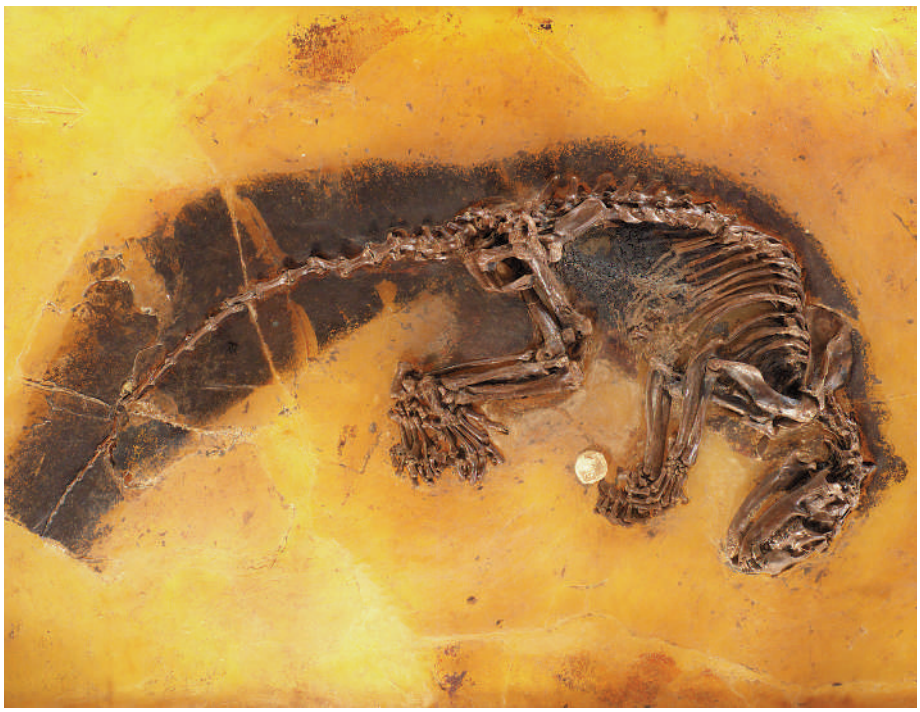
One of the charms of this book is the abundance of figures, with many attractive reconstructions of Eocene mammals. Some of these have been made possible by the unusual preservation of complete mammalian skeletons from the Middle Eocene in Messel, Germany. Rose had an opportunity to collaborate with German colleagues when describing some of these mammals. As well as line drawings and black-and-white photographs, there are eight full-page colour plates with photographs of the best-preserved fossils and reconstructions.

Although the bulk of the book is devoted to this mammalian explosion, Rose starts at the beginning of the mammals' tale. Two introductory chapters are followed by four on the origin of mammals and their early evolution in the Mesozoic.

The author no doubt faced a problem when dividing the rest of the book into chapters. Details about the roots of many groups are hazy, and the earliest representatives of most orders are unknown, so it would be impossible to restrict the contents of chapters to strictly monophyletic units. Nowadays, taxonomists have to deal with two different taxonomies: a traditional one, based on morphology, and a modern one, based on molecular analysis. Each chapter begins with an introduction describing our changing views on the relationships of the animals to be discussed. These introductions show the perspective and knowledge that Rose uses when examining the data. When available, he provides two alternative cladograms showing the relationships among the studied forms, one based on morphology, the other on molecular studies. His preference is decidedly morphological, as is evident from his attitude to the clade Afrotheria.

The notion of Afrotheria arose around 1990, based on DNA analyses demonstrating that six living orders of African placental mammals form a monophyletic clade. Although Rose often mentions the Afrotheria, he states: "No morphological evidence supporting Afrotheria has been found." However, the Cretaceous representatives of modern mammalian orders are unknown, and the lack of an early fossil record for most groups may be responsible for any morphological gaps between them. Among the groups assigned to Afrotheria, Rose, on the basis of morphology, accepts the monophyly of Hyracoidea, Elephantoidea and Sirenia. But, against the molecular evidence, he assigns Tenrecidae and Chrysochloridae not to Afrosoricida but to Insectivora; Macroscelidea to the relatives of the Anagalida; and finally Tubulidentata (not described because of a lack of fossils) to the archaic Ungulata.

Each chapter contains a systematic table using linnaean taxa and a hierarchy of systematic units following Malcolm McKenna and



FORSCHUNGINST. SENCKENBERG, FRANKFURT

Fossil skeletons of Middle Eocene mammals such as *Kopidodon* have been found in Messel, Germany.

Susan Bell's book *Classification of Mammals* (Columbia University Press, 1997). But the PhyloCode (www.ohio.edu/phylocode) — a set of rules governing phylogenetic nomenclature that in recent years has gained many proponents among zoologists and palaeontologists — is not mentioned in the book. In the final chapter, Rose provides a summary of the idea of a sudden explosion of mammals after the Cretaceous–Tertiary boundary.

This beautiful and thorough book will be an essential tool for all those who work on fossil and extant mammals, and for both advanced undergraduate and graduate students. It is a 'must buy' for palaeontological libraries. ■

Zofia Kielan-Jaworowska is at the Institute of Paleobiology, Polish Academy of Sciences, Warsaw, Poland.

Mysteries of female anatomy

Secrets of Women: Gender, Generation, and the Origins of Human Dissection
by Katharine Park
Zone Books: 2006. 300 pp. $36.95

Rina Knoeff

Although the history of dissection in general has received increasing attention, the importance of anatomizing female bodies has hitherto been neglected. Katharine Park's *Secrets of Women* is a first attempt to fill the gap. She relates a fascinating history of women on the dissection table. Central to her book are the almost unbelievable stories of an abbess whose heart contained a cross, a crown of thorns, nails and a lance, items well known as symbols of the Passion of Christ; the investigation of the body of a lactating virgin prophetess; the search for the cause of death of patrician wives and mothers; and the dissected body of an executed female criminal. The accounts are bizarre yet captivating, as they portray the opening of the body not only as a medical procedure, but also as a cultural practice embedded in religious and social life. Park also exposes the myths that in the wake of Pope Boniface VIII's bull Detestande feritatis in 1299, the opening of bodies was surrounded by a cultural taboo, and also that anatomical practice was primarily a male practice studying male bodies.

BRITISH LIBRARY, LONDON/BRIDGEMAN ART LIBRARY

The dissection of female bodies had religious and familial significance in the fifteenth century.

Park argues that from the mid-thirteenth to the mid-sixteenth century, Italians understood their bodies both in terms of religion and in relation to family and kinship, and that this particular way of viewing the body shaped dissection practices. Male uncertainty regarding fatherhood (were they really the fathers of their children?), the existence of female-only knowledge regarding sexuality and generation (mainly related to inordinate sexual appetite, abortion practices, the exercise of sexual autonomy and the exertion of power over men), as well as the functioning of the reproductive system inside the female body, culminated in an excessive interest in what came to be known as 'the secrets of women'. As a result, Park argues, the uterus (and hence the inside of the female body) became a focal point of anatomical research in Italian patriarchal society.

Related to the opening of mothers' bodies, dissection practices concentrated on holy women. Not only did women manifest religious experiences in a particular corporeal way, but these experiences were also described in physiological terms of pregnancy. It was, for instance, believed that Christ's presence in the heart would create new life in the same way as the father's seed in the mother's uterus. According to Park, the themes of holiness, generation and female corporeality together contributed importantly to the history of dissection.

At times, however, Park's interpretation is merely suggestive and the historical proof wafer-thin. For instance, her argument that holy women generate relics much as mothers generate children is well argued. But her suggestion that the dissection of holy women in order to extract relics can be viewed as a religious variant of a caesarean section, in which a baby was removed from a dead mother's body so it could be baptized, seems far-fetched.

In a similar vein, Park's understanding of religion is sometimes rather naive. Her interpretation of Christianity as exclusively Roman Catholic is understandable given the Italian context she describes, but it leaves no room for a pluralist view of religion. For Park, depictions that lack characteristically Roman Catholic symbols don't seem to count. For example, she interprets the frontispiece of Andreas Vesalius' book on human anatomy *De humani corporis fabrica* as secular, simply because there are no Catholic icons. Yet, particularly given Vesalius' Protestant upbringing, it can, and perhaps should, be viewed as a Protestant depiction (as other historians of medicine have done before).

Nevertheless, Park's book will undoubtedly prove to be an important contribution to the history of anatomy. For the first time it extensively discusses the history of anatomy from the viewpoint of the corpse and, because of its particular focus on women's bodies, it will radically change the way we think about the (male) history of the anatomized body. As such, the book is a 'must read' for anyone working on the history of pre-modern medicine. But because of its lucid style and fascinating argument, it is also surprisingly accessible, and I recommend it to anyone interested in the history of anatomy. ■

Rina Knoeff is in the Faculty of Arts, University of Leiden, 2300 RA Leiden, the Netherlands.

The structure of consciousness

Subjective awareness may depend on neural networks in the brain supporting complex wiring schemes and dynamic patterns of activity.

György Buzsáki

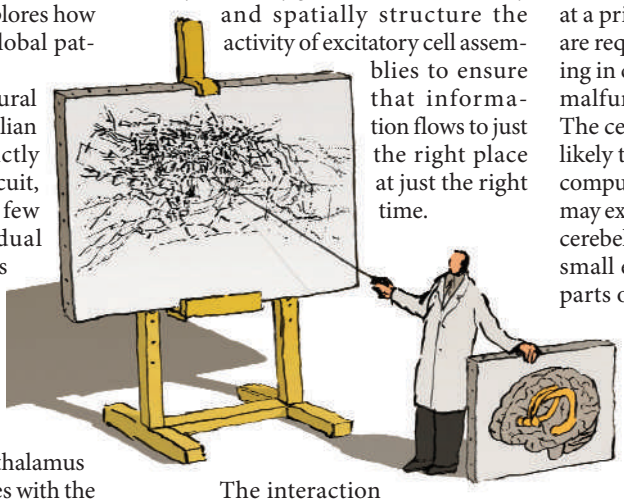
Perhaps nowhere is the truism 'structure defines function' more appropriate than for the brain. The architecture of different brain regions determines the kinds of computations that can be carried out, and may dictate whether a particular region can support subjective awareness. Also, the degree of architectural complexity may determine susceptibility to neurological and psychiatric diseases — complex architectural schemes being more prone to disruption than simpler ones. Understanding how such structure–function relationships govern brain operations requires a systems-level approach that explores how local computation relates to global patterns of neural activity.

At least three basic architectural schemes are present in mammalian brains. The simplest uses strictly local wiring. In this kind of circuit, typified by the cerebellum, a few neuronal types form individual 'modules' that may be repeated as necessary. Because interaction between modules is restricted to neighbours, it is massively parallel in nature. In different species, the size of locally organized brain structures — including the basal ganglia, thalamus and cerebellum — roughly scales with the number of modules they contain.

An entirely different type of network uses random connections, with a more or less equal probability of connecting local, intermediate or distant neurons. A unique example of such a random connectionist scheme is the recurrent excitatory circuit of the hippocampal CA3 region.

The third architectural scheme, exemplified by the neocortex, combines local modularity with more random, long-range connectivity. This complex wiring scheme shares many properties with 'small-world' or 'scale-free' networks. The advantage of this arrangement is that the number of intermediate steps between any two neurons — the synaptic path length — can remain relatively constant when network size is scaled up, because even a small fraction of long-range connections can dramatically reduce the average path length. Although intermediate and long-range interconnections demand resources and space, they are critical for globally distributing the results of local computations throughout the entire cerebral cortex.

I propose that the distinct network architectures translate into unique functional consequences. In cortical networks, a dynamic balance between excitation and inhibition gives rise to an array of network oscillations involving neuronal populations of varying sizes. This self-organized, or so called 'spontaneous' activity is the most striking and yet perhaps least appreciated feature of the cerebral cortex. Without inhibition, excitatory activity caused by any one stimulus would ripple across the entire neuronal network and a confused jumble of overlapping signals would result. Inhibitory interneurons and the rhythms they generate can temporally and spatially structure the activity of excitatory cell assemblies to ensure that information flows to just the right place at just the right time.



J. KAPUSTA/IMAGES.COM

The interaction and interference of multiple brain rhythms often gives rise to the appearance of 'noise' in an electroencephalogram. This noise is the most complex type known to physics and reflects a metastable state between the predictable behaviour of oscillators and the unpredictability of chaos. Neural firing patterns are thus controlled not only by the external sensory environment but also by the internally generated and perpetually changing state of cortical networks. Because local computation can be sensed by large parts of the cortex through long-range connections, and is also modified by this background 'noise', the term 'local–global computation' best captures the nature of cortical operations. A special case is the hippocampus whose highly recursive connection matrix is thought to function as a large 'autoassociator', allowing the reconstruction of entire episodes from remembered fragments.

I suggest that the local–global wiring of the cerebral cortex and the perpetual, self-organized complex dynamics it supports are necessary ingredients for subjective experiences. Environmental inputs can be seen as perturbations of the ongoing spontaneous activity. If they manage to perturb ongoing activity for a sufficiently long time in a big enough population of neurons, their effect will be noticed; that is, we will become conscious of them. In contrast, the locally organized cerebellar cortex, used largely for sensorimotor integration, does not give rise to self-generated or spontaneous activity, and its response to input remains local and non-persistent. Importantly, we generate no subjective record of such local computations.

Complex neuronal networks are a useful product of brain evolution but come at a price. Greater resources and volume are required to sustain long-distance wiring in complex networks, and the risks of malfunction increase with complexity. The cerebellar-type local organization is likely to be robust because errors in local computation are not disseminated, which may explain why functional diseases of the cerebellum are relatively rare. In contrast, small errors can propagate across large parts of complex networks and perhaps even be magnified by the rhythm-regenerating properties of the cerebral cortex. Timing errors present particularly difficult problems in complex networks, because of limits to how much information can be conveyed through restricted numbers of long-range conduits. Not surprisingly, diseases of the cerebral cortex are manifold — including epilepsies, Alzheimer's and schizophrenia.

One of the greatest challenges left for systems neuroscience is to understand the normal and dysfunctional operations of the cerebral cortex by relating local and global patterns of activity at timescales relevant for behaviour. This will require monitoring methods that can survey a sufficiently large neuronal space at the resolution of single neurons, and computational solutions that can make sense of complex interactions. ■

György Buzsáki is at the Center for Molecular and Behavioral Neuroscience, Rutgers University, 197 University Avenue, Newark, New Jersey 07102, USA.

FURTHER READING
Tononi, G., Sporns, O. & Edelman, G. M. *Proc. Natl Acad. Sci. USA* **91,** 5033–5037 (1994).
Buzsáki, G. *Rhythms of the Brain* (Oxford Univ. Press, 2006).

For other essays in this series, see http://nature.com/nature/focus/arts/connections/index.html

PLANT TAXONOMY

The love of plants

Carl Linnaeus's use of erotic language to describe plants ultimately helped him to recruit a global network of specimen collectors.

Staffan Müller-Wille

In August 1749, Pehr Kalm, a medical student from Finland, travelled through Quebec on behalf of the Royal Swedish Academy of Sciences. The members of the academy, among them the botanist Carl Linnaeus, had sent him there in the hope that he would find useful plants to boost the Swedish economy. While in Quebec, Kalm observed a startling behaviour among the locals. All over the province, he wrote in his travel journal, "high officers, fortress commanders, priests, Jesuits, soldiers, tradesmen, and even the nobler sex, were to be seen running about the forests gathering plants as though they were trained by Doctor Linnaeus".

Kalm himself had taken part in Linnaeus's famed excursions into the surroundings of Uppsala University — colourful events involving hundreds of academics, students and lay-people alike — where he had learned how to collect and correctly identify plants. It must have been astounding for Kalm to see that this custom had also been taken up in what was then one of the remotest corners of the world.

Kalm's observation shows the degree to which Linnaeus's fame had spread around the globe in the mid-eighteenth century. The secret of this success lay in an ingenious tool that Linnaeus had designed when teaching botany at Uppsala University. Known as the sexual system of plants, it was a simple classification system whereby the number and position of the reproductive organs in the flower were used to determine the class and order to which a plant belonged. Linnaeus described this system in the *Systema Naturae*, a major work published in 1735 in which he laid out his ideas about classifying the animal, plant and mineral kingdoms. There, Linnaeus referred to flowers as 'beds' and stamina and pistils as 'husbands' and 'wives'. Indeed, the sexual system comprised a veritable catalogue of eighteenth-century sexual relations, from monogamy and polygyny, to homosexuality, miscegenation and incest. Using the sexual system to classify plants, especially in company, would not have been an entirely innocent affair. For many, it must have turned botany into a highly amusing, if not stimulating, pastime.

The 'sexual system' sparked interest in plants.

BRIDGEMAN ART LIBRARY

But the success of the sexual system came at a price. Some, such as the botanist Johann Georg Siegesbeck, accused Linnaeus of subjecting students to immoral influences. More serious attacks came from naturalists such as the Comte de Buffon, who wanted more from taxonomy than mere distinctions. For all the amusement it may have provided, ultimately the sexual system said little about plants and the relationships between species. During his lifetime, Linnaeus saw himself portrayed as a dry scholastic, whose interest lay in diagnostic definitions, not in biological questions.

If the sexual system had been his only achievement, Linnaeus would probably have whole-heartedly agreed with this verdict. He used to emphasize that the sexual system was artificial — nothing but a diagnostic tool for beginners — and was clearly aware of the fact that organisms were interrelated by a complex web of associations that could not be unravelled by simple definitions. "The true beginning and end of botany," he wrote in 1751 in the *Philosophia Botanica*, "is the natural system", a system in which "all plants exhibit mutual affinities, as territories on a geographical map".

The natural system was supposed to pay attention not to a few selected characters, but to the overall morphology of flower and fruit. Working out the 'natural system' of plants required empirical work on a huge scale. Plant specimens had to be procured from all over the world and meticulously collated, described and compared. In 1753, Linnaeus published *Species Plantarum*, a book listing the names and definitions of more than 5,000 plant species. While preparing the book for publication, Linnaeus complained to his close friend Abraham Bäck, a physician in Stockholm, that he felt like a hen, laying an egg a day.

But as well as the dedication needed to work through thousands of specimens, piecing together the natural system required exceptional abilities in recruiting correspondents and travellers. Linnaeus was a genius in networking. In one of his autobiographies, he described his vast network of correspondents as an army, with Europe's leading botanists serving as generals, and lay-people as far away as Istanbul and North America serving as footmen. It was in establishing this network that the sexual system played a crucial role.

The popularity of this system created a large number of adepts in linnaean botany all over the world, who would send him specimens and seeds, following the conventions he had introduced with respect to naming and describing new species. In this way, the sexual system served as a recruitment tool to engage the public in a research project of a grand scale: the quest for the natural system. And it worked spectacularly. In 1763, Linnaeus even named a new species *Quassia*, in honour of an African-American slave from Suriname who had communicated specimens of that species to him.

Since Linnaeus put the natural system on the map, plant taxonomy has undergone numerous technological and conceptual revolutions. But in 2007, the tercentenary of Linnaeus's birth, the quest for the natural system continues. Today, plant taxonomists use sophisticated computational and molecular tools to uncover the relationships of descent that underwrite the natural system. Plant taxonomy may have lost some of the charm it held for the public in the eighteenth century, but it is still one of the most vibrant research areas in the life sciences. ■

Staffan Müller-Wille is at the ESRC Centre for Genomics in Society, University of Exeter, Exeter EX4 4PJ, UK.

FURTHER READING
Blunt, W. *Linnaeus: The Compleat Naturalist* 2nd edn (Frances Lincoln, London, 2004).
Koerner, L. *Linnaeus: Nature and Nation* (Harvard Univ. Press, Cambridge, MA, 1999).
Natural History Museum, London *The Linnaean Plant Name Typification Project* www.nhm.ac.uk/research-curation/projects/linnaean-typification/index.html.

CONCEPTS

NEWS & VIEWS

BOTANY

New home for tiny aquatics

Else Marie Friis and Peter Crane

A shake-up of current thinking about the evolution of the angiosperms — the flowering plants — is a consequence of the relocation of a hitherto obscure branch on the angiosperm evolutionary tree.

Many new constellations in the angiosperm evolutionary tree have come to light as relationships within the flowering plants have been explored using molecular analyses[1,2]. Some of these changes were not unexpected. But others have resulted in fundamental reinterpretations of angiosperm evolution[3,4].

On page 312 of this issue, Saarela *et al.*[5] report one of the most striking realignments required so far. They find that the diminutive, moss-like, aquatic plants *Hydatella* and *Trithuria*, which are members of the family Hydatellaceae and were thought to be monocots close to grasses, are actually the closest living relatives of water-lilies and their allies (Fig. 1). This new-found position for the Hydatellaceae, near the point at which water-lilies (order Nymphaeales) diverged from other flowering plants, could scarcely be more attention-grabbing. On the evidence of Saarela and colleagues' analyses[5], and given other generally accepted relationships at the base of the angiosperm evolutionary tree[2,3], only a single angiosperm species, *Amborella trichopoda*, diverged from other flowering plants below this point.

But is this new discovery simply a minor matter of phylogenetic tidying up? After all, the Hydatellaceae are a small family; their position has long been uncertain[6]; and the unexpected is bound to crop up when relationships in a group the size of flowering plants (some 350,000–400,000 extant species) are looked at more carefully.

At one level, as Saarela *et al.*[5] point out, repositioning of the Hydatellaceae conforms to, rather than overthrows, current ideas of relationships among the extant representatives of early angiosperm lineages. These ideas have come to look increasingly secure, as more genes and more plants have been incorporated into molecular phylogenetic analyses. The Hydatellaceae associate with the Nymphaeales based on the new molecular data, and the relationship seems to be well corroborated. Many of the morphological characteristics of the Hydatellaceae, at least in so far as they are known, also make more sense in the context of the Nymphaeales than in the context of grasses and their relatives. And it is convenient that the Hydatellaceae, as a family of aquatic plants, link to the Nymphaeales rather than to the other, mainly woody, plants that make up the root of the angiosperm evolutionary tree. In this position they do not further complicate ideas about the evolution of life in an aquatic habitat.

For other traits, however, repositioning the Hydatellaceae raises questions that add to an already long list of unresolved issues in early angiosperm evolution, particularly with regard to features that are perhaps too easily interpreted as 'reduced', 'lost' or 'absent'. At some point it may become more straightforward to infer that some early angiosperms never had certain features, rather than had them and then lost them.

For example, does the seemingly simple floral morphology of the Hydatellaceae reflect

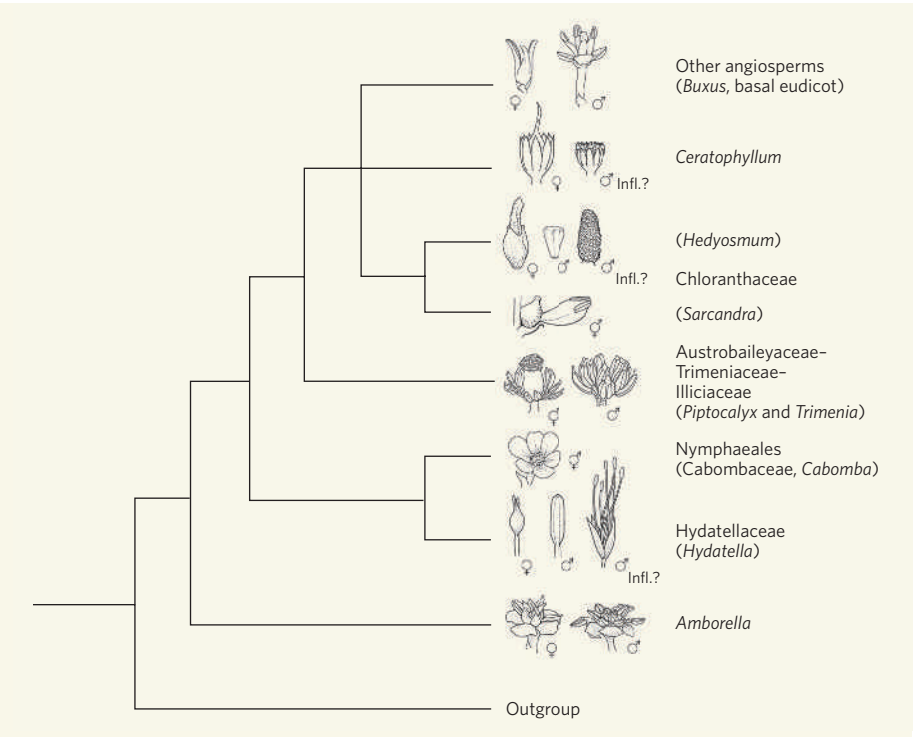


Figure 1 | A new branch near the root of the angiosperm evolutionary tree. *Amborella* remains the sister group to all other flowering plants, but the Hydatellaceae join the Nymphaeales (water-lilies) as the next diverging subsidiary branch of the tree. This new alignment[5] raises questions about the origin of classic bisexual flowers in the Nymphaeales and other angiosperms. The unisexual flowers of Hydatellaceae — with male and female flowers consisting of a single organ only (stamen and carpel, respectively), aggregated in dense unisexual or bisexual inflorescences — contrast with the solitary, bisexual flowers of the Cabombaceae, represented here by *Cabomba*, which were previously regarded as the basic condition in the Nymphaeales. More generally, simple and often unisexual flowers occur among several early-diverging angiosperm lineages, for example in the Trimeniaceae (*Piptocalyx, Trimenia*) and Chloranthaceae (*Sarcandra*, *Hedyosmum*), as well as early-diverging eudicots (*Buxus*). But whether this simplicity reflects the basic phylogenetic condition or ecological adaptation is an open question. Flowers of *Amborella* are also small, and functionally unisexual, but female flowers have staminodes, indicating a basic bisexual organization. In this situation, interpretations of character evolution depend on the position of the root of the phylogenetic tree, and may also be sensitive to the addition of new taxonomic groups. Infl, inflorescence (flower cluster). (Drawings by P. von Knorring.)

D. STEVENSON, NY BOTANIC GARDEN

Figure 2 | *Trithuria* — not previously on the radar screen of most specialists.

'reduction', or might it represent an early 'pre-floral' stage of angiosperm evolution in which the classic bisexual flower, with its whorls of three or more different organs, was not yet fully formed? Flowers of the Hydatellaceae — along with those of other early-diverging angiosperm lineages (for example, fossil and living *Hedyosmum*[7] and other members of the Chloranthaceae) — could hardly be more simple (Fig. 1). It is possible, in the case of the Hydatellaceae, as well as *Ceratophyllum* and perhaps the fossil *Archaefructus*[8,9], that the simple flowers reflect loss of floral parts associated with life in a submerged habitat, as has happened in other aquatic angiosperm lineages. But here, near the base of the angiosperm phylogenetic tree, it is difficult to be sure whether floral simplicity signals ecological adaptation or the basic phylogenetic condition. If it is the latter, it would have profound implications for ideas on the early evolution of the classic angiosperm flower.

It will take some time to digest all the implications of suddenly introducing a new plant family into discussions of early angiosperm evolution. *Hydatella* and *Trithuria* (Fig. 2) have not been on the radar screen of most specialists working on the subject, and there are many pieces of this puzzle still to work out. Certainly, repositioning the Hydatellaceae as the closest living relatives to the Nymphaeales will dramatically influence ideas about the early evolution of water-lilies and their allies, and it will modify important details of character evolution at the base of the angiosperm tree. Whether it will also affect the rooting of the angiosperm tree as a whole, perhaps in ways that would displace *Amborella* to a less prominent position, remains to be seen. Identification of angiosperm precursors in the fossil record would probably have an even greater impact.

It is also relevant that several species of Hydatellaceae have been described only in the past 25 years. It seems likely that more will come to light, and there may also be revelations among the several hundred plant species that are described as new to science every year. Examples such as the Hydatellaceae, and the Australian conifer genus *Wollemia*[10] discovered a little over a decade ago as a 'living fossil', illustrate the inadequacy of our knowledge of plant diversity at a time when so much is being destroyed or is under serious threat.

Hydatella and *Trithuria* will inevitably be the subject of detailed investigation in the coming years. But whatever the outcomes of these studies, the radical realignment discovered by Saarela *et al.*[5] should remind us not to become too comfortable with the current picture of early angiosperm relationships, and especially with the details of character evolution that they imply. There will be more surprises as new plants are added to the mix. They will come not just from our gradually improving knowledge of living plants, but more especially from our exploration of the riches of the plant fossil record — both for early angiosperms and for their elusive relatives among other seed plants. ■

Else Marie Friis is in the Department of Palaeobotany, The Swedish Museum of Natural History, Box 50007, 104 05 Stockholm, Sweden. Peter Crane is in the Department of the Geophysical Sciences, The University of Chicago, Chicago, Illinois 60637, USA.
e-mails: elsemarie.friis@nrm.se
pcrane@geosci.uchicago.edu

1. Chase, M. W. *et al. Ann. Missouri Bot. Gard.* **80,** 528–580 (1993).
2. Soltis, D. E., Soltis, P. S., Endress, P. K. & Chase, M. W. *Phylogeny and Evolution of Angiosperms* (Sinauer, Sunderland, MA, 2005).
3. Qiu, Y.-L. *et al. Nature* **402,** 404–407 (1999).
4. Endress, P. K. & Igersheim, A. *Int. J. Plant Sci.* **161,** S211–S223 (2000).
5. Saarela, J. M. *et al. Nature* **446,** 312–315 (2007).
6. Dahlgren, R. M. T., Clifford, H. T. & Yeo, P. F. *The Families of the Monocotyledons: Structure, Evolution, and Taxonomy* (Springer, Berlin, 1985).
7. Friis, E. M., Pedersen, K. R. & Crane, P. R. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **232,** 251–293 (2006).
8. Sun, G. *et al. Science* **296,** 899–904 (2002).
9. Friis, E. M., Doyle, J. A., Endress, P. K. & Leng, Q. *Trends Plant Sci.* **8,** 369–373 (2003).
10. Jones, W. G. *et al. Telopea* **6,** 173–176 (1995).

PHYSICAL CHEMISTRY

The peripatetic proton

James T. Hynes

The way in which protons are transferred between acids and bases has been known in general terms for decades. But the details of the process are complex, and only now is the full proton itinerary becoming clear.

Proton transfer is a headline player in many arenas of chemistry and biology. To name just a few, the process lies at the heart of the chemistry of acids and bases in solution, the workings of enzymes, and transport mechanisms in biological membranes and photosystems[1,2]. A broad-brush picture of how protons are transferred between acids and bases in aqueous solution was painted in the 1950s and 1960s, in classic work by Eigen[3] and Weller[4]. Writing in *Angewandte Chemie International Edition*, Mohammed *et al.*[5] enlarge that picture and add more detail. They build on earlier, related experiments of their own and those of others to show how modern, ultra-rapid techniques can illuminate the myriad, sometimes indirect, molecular pathways between acid and base that a proton can follow.

An acid is a substance that likes to donate protons; a base is a substance that is inclined to accept them. Mohammed *et al.* used laser excitation to 'trigger' the departure of a proton from the light-sensitive acidic molecule pyranine. This acid can be written ROH, where R is an organic group and OH a hydroxyl group. As a proton is simply a hydrogen atom stripped of its electron, its progress can be tracked by identifying, through spectroscopy, where the 'H' appears in the chemical products of the subsequent reactions.

Using ultra-fast infrared vibrational spectroscopy as a structural probe, the authors could in their experiments reconstruct the dynamics of the proton's entire voyage — from its departure from ROH, leaving behind the base RO^-, to its ultimate destination, a negatively charged base molecule, denoted B^-. This base molecule is actually a trichloroacetate anion, $^-OOCCCl_3$, to which the proton attaches itself to produce a carboxylic acid, or BH in our notation. Different concentrations of the base, from 1 to 3 mol l^{-1}, helped the authors to unravel the exact proton-transfer dynamics.

The authors' first significant finding[5] is that infrared bands characteristic of absorption by RO^- are observed on a timescale of less than 150 femtoseconds (1 femtosecond is 10^{-15} s) after the initial laser excitation. These are not, however, immediately accompanied by the appearance of any absorption band corresponding to BH. Indeed, at a B^- concentration of 2 mol l^{-1}, protonation to form BH is incomplete even after 700 picoseconds — more than 1,000 times longer than it takes for ROH to be deprotonated.

So where is the proton in the meantime? Making intermediate stops at water molecules, say Mohammed and colleagues. They observe a signal indicating the presence of hydronium ions, H_3O^+, on the same short timescale as the RO^- absorption band. They argue that this

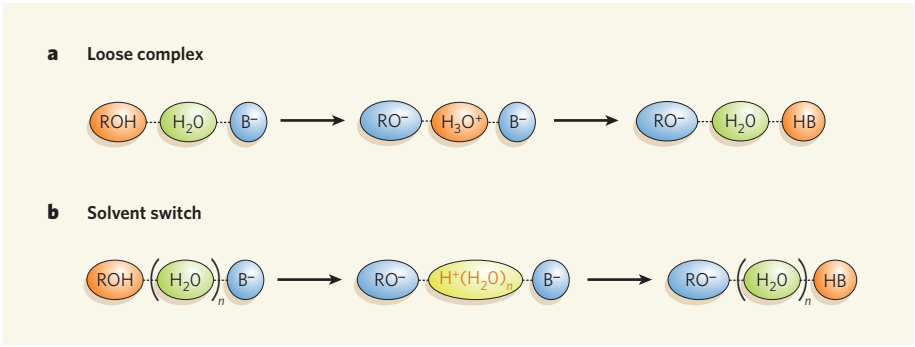


Figure 1 | Water stop. In their detailed spectroscopic investigations of proton (H^+) transfer between a laser-excited acid (ROH) and a base (B^-) in aqueous solution, Mohammed *et al.*[5] observe two separate signal components. **a**, An ultra-fast rise in absorption bands corresponding to the deprotonated form of the acid, RO^-, and to the hydronium ion, H_3O^+, indicates the formation of a simple, loosely hydrogen-bound complex between an acid, a base and an intervening water molecule, as a short-lived precursor to the full transfer of the proton to the base, forming a carboxylic acid (here, HB). This loose-complex component accounts for only about 18% of the laser-excited ROH. **b**, The slower rise of the absorption band corresponding to HB, however, indicates that other, more complex processes are also at work: in most cases, the itinerant proton makes stops at several intermediate water molecules $(H_2O)_n$, a process known as a solvent switch. Mohammed *et al.*[5] used the deuterated forms of acid and water in their experiments, but the principles remain the same.

signal arises almost exclusively when a single water molecule, H_2O, intervenes in a loose complex of ROH and B^-. This intervention allows it to gain a proton from ROH, thus generating a 'loose' hydrogen-bonded complex, $RO^-\ldots H_3O^+\ldots B^-$ (Fig. 1a).

Such a loose complex would arise only about 18% of the time at a B^- concentration of 2 mol l^{-1}. That percentage is consistent with the observation that the initial ultra-fast rise of the RO^- band amounts to some 18% of the total laser-excited ROH present, and is followed by a much slower form of exponential rise. Evidently, ROH is losing protons through more complicated, time-consuming pathways than just the simple confinement of the proton in a loose complex.

An insight into what these alternative pathways might be is provided by the authors' second finding: that the rate of disappearance of the H_3O^+ band does not match that of the rise of the BH band. It seems that the proton is not transferred directly to the base B^- in the loose complex, but follows an indirect route to its final destination, shuttling stepwise over one or more additional water molecules in its quest for the base. This process is termed a solvent switch (Fig. 1b).

The involvement of more than one water molecule in the proton's route to the base is consistent with the slower exponential rise in the RO^- band already mentioned. But the dominance of the solvent-switch mechanism contrasts with the authors' earlier results in a study involving the same acid but a stronger base with a greater voracity for protons[6]. In that case, there is only one sequence of stepping-stones for the proton to cross, from the acid to a loose complex to the base (Fig. 1a) — a mode of locomotion known as the Grotthus hopping mechanism[1,2,7].

Even with the solvent-switch mechanism dominant, the proton is only ever directly observed localized on a single water molecule in the loose complex. As the authors explain, the spectral tracking of such a meandering proton — possibly also hopping in a Grotthus fashion — is a ticklish problem. Detailed support for the solvent-switch mechanism is instead provided by a kinetic model that involves proton transfer through various numbers of water molecules intervening between the acid and the base, together with the rates of interconversion between the different reactants and products. The wealth of rate constants extracted not only buttresses the authors' interpretation, but also allows an assessment of the relative likelihood of a proton being transferred in a complex involving a certain number of water molecules. This can then be compared with the probability that a water molecule will be ejected from the complex instead, so that the transfer happens with one fewer molecule present.

This last point brings us back to the classic ideas of Eigen[3] and Weller[4], established with lower-resolution time and much less sensitive tools for assessing structure. These researchers constructed a picture in which an acid and a base would diffuse to within a certain distance of each other, at which point the proton would transfer in some unknown fashion over an imprecisely specified number of water molecules. The work of Mohammed and colleagues[5], together with other proton-transfer studies[1,2], represents significant progress on the path towards a full, molecular-level characterization of the processes involved. But this ultimate goal has not yet been met. We can still look forward, for example, to the direct observation of the shuttling proton in all its assorted pathways, the classification of these pathways according to the characteristics of the acid and the base, and the elucidation of the quantum-mechanical features of the proton's movement[7]. ■

James T. Hynes is in the Department of Chemistry and Biochemistry, University of

50 YEARS AGO

In the fifth of its reports, "Federal Funds for Science", the National Science Foundation analyses the U.S. Federal research and development budget for the fiscal years 1955, 1956 and 1957... The total U.S. Federal Government expenditure on research and development is estimated at about 2,300 million, 2,400 million and 2,700 million dollars in 1955, 1956 and 1957, respectively... In scientific fields, the physical sciences (including engineering sciences) continued to claim the major share, the proportion during 1955–57 being 87 per cent compared with 11 per cent for the biological sciences and 2 per cent for the social sciences.

From *Nature* 16 March 1957.

100 YEARS AGO

Dr. Rivers has re-discovered the Todas. This curious little nation, long known to us as an isolated social abnormality, in which the dairy industry takes the place of religion and matrimonial safety is found in a plurality of husbands, now appears to be both much more and much less than this... The chief regulations of the marriage system are in brief :— prohibition of intermarriage between the two "castes" *Tartharol* and *Teivaliol*; exogamy among the clans which compose these "castes"; certain kinship prohibitions; polyandry, the typical form of marital association, the extra husbands being generally the brothers of the husband proper; polygyny, now on the increase, either in the ordinary form, or two men having two wives in common... The economic sources of religion are more clearly laid bare in the full description of the dairy-religion of the Todas... To quote Dr. Rivers :— "The sacred animals are attended by men especially set apart who form the Toda priesthood, and the milk of the sacred animals is churned in dairies which may be regarded as the Toda temples, and are so regarded by the people themselves."

From *Nature* 14 March 1907.

50 & 100 YEARS AGO

Colorado, Boulder, Colorado 80309-0215, USA, and the Département de Chimie, Ecole Normale Supérieure, Paris, France.
e-mail: hynes@chimie.ens.fr

1. Hynes, J. T., Klinman, J. P., Limbach, H.-H. & Schowen, R. L. (eds) *Hydrogen-Transfer Reactions* Vols 1–4 (Wiley-VCH, Weinheim, 2007).
2. Elsaesser, T. & Bakker, H. J. (eds) *Ultrafast Hydrogen Bonding Dynamics and Proton Transfer Processes in the Condensed Phase* (Kluwer, Dordrecht, 2002).
3. Eigen, M. *Angew. Chem. Int. Edn* **2,** 285–318 (1964).
4. Weller, A. *Prog. React. Kinet.* **1,** 187–213 (1961).
5. Mohammed, O. F., Pines, D., Nibbering, E. T. J. & Pines, E. *Angew. Chem. Int. Edn* **46,** 1458–1461 (2007).
6. Mohammed, O. F., Pines, D., Dreyer, E., Pines, E. & Nibbering, E. T. J. *Science* **310,** 83–86 (2005).
7. Ando, K. & Hynes, J. T. *J. Phys. Chem. A* **103,** 10398–10408 (1999).

SOLAR SYSTEM

Portrait of a suburban family

Alessandro Morbidelli

The first 'collisional family' has been spotted among objects in the Kuiper belt, which lies on the outskirts of the Solar System. The identification could provide useful constraints on the outer Solar System's history.

Collisional families — clusters of Solar System bodies formed when their parent body breaks up on being hit by a projectile — are familiar to astronomers. About 35 of them are known in the asteroid belt[1], the ring of small 'planetesimal' bodies confined between the orbits of Mars and Jupiter. On page 294 of this issue[2], Brown and colleagues report the first identification of a collisional family in the Kuiper belt, another region of planetesimals that lies in the suburbs of the Solar System beyond the orbit of Neptune. This identification, and the inferences that can be drawn from it, are especially significant because the Kuiper belt is believed to be the relic of the huge primordial disk of planetesimals from which the cores of the giant planets formed.

For a collisional family to be observed from Earth, its parent must have been hit with enough energy to disperse a large number of large fragments. Families are identified as groups of bodies whose orbital characteristics are sufficiently similar for them to be statistically relevant[3]. The families in the asteroid belt formed from parents with diameters ranging from 10 to 450 kilometres, which were typically hit by projectiles less than half their size[1,4].

Going by the size of its members, Brown and colleagues' family[2] has no analogue in the asteroid belt. The largest member, now formally designated (136108) 2003 EL_{61}, is one of the largest objects in the Kuiper belt: its shape and mean diameter of about 1,500 km make it a contender as a 'dwarf planet'. The five other members of the family identified are tightly clustered, and have diameters of between 10 and 400 km. They have similar physical properties: a spectrum with deep absorption bands characteristic of the presence of surface water ice, and a grey colour. No other object so far discovered in the Kuiper belt has both these properties. Furthermore, 2003 EL_{61} has several oddities: it is an anomalously fast rotator; its bulk density is estimated to be larger than that of any other Kuiper belt object; and it has two moons. These features had already seemed to indicate that 2003 EL_{61} was originally a body with a high-density rocky core and a low-density icy mantle. It must then have suffered a catastrophic collision — with a projectile around 1,000 km in diameter — that spun it up and ejected a large fraction of its icy mantle into space[5] (Fig. 1). Taken together, all these aspects make Brown and colleagues' claim for a genetic relationship between these bodies compelling.

The discovery of such a family is important for at least three reasons. First, it gives us the first opportunity to test our understanding of collisional physics at the size and energy scales characteristic of collisions between dwarf planets. The formation of the Pluto–Charon binary system and of Earth's Moon have both been modelled successfully as the result of giant collisions between dwarf or fully grown planets[6,7]. The new family presents a much harder challenge to modellers because, unlike the Pluto–Charon and Earth–Moon cases, the simulations will also have to reproduce the existence of the collection of many large fragments orbiting the Sun, and their dispersion.

The second reason that this family arouses our interest is that it must have been formed at a specific time in the history of the Solar System. Estimates of the current Kuiper belt population[8] indicate that there should be about 10 objects with a diameter about as big as that of the family's progenitor, and about 50 as large as the projectile that hit it. Taking those figures at face value, and given the orbital geometries in the Kuiper belt, the probability, over the age of the Solar System, of the right collision occurring to form the family is only about 1 in 2,000.

We have many indications that the Kuiper belt was initially much more massive than it is today[9]; presumably, therefore, the collision must have happened in the distant past, when the belt was far more crowded with large objects. But the situation is more complex than this. The orbits of the planetesimals in the primordial disk should all have been circular and in the same plane. But the orbits of the bodies in the Kuiper belt exhibit a variety of eccentricities — a measure of an orbit's deviation from a perfect circle — and inclinations. These range up to large values: 2003 EL_{61} itself has an orbital eccentricity of about 20%, and an inclination of almost 30°. It is generally believed that this 'sculpting' of the orbits into a complex structure occurred at more or less the same time that most of the primordial mass was removed. The fact that Brown and colleagues' family retained similar orbital characteristics implies that it was formed after the sculpting — because otherwise the family would have been dispersed in that event — but before the Kuiper belt population became too small for enough projectiles to have been around to bring about a collision. The need for a time window during which these two constraints are fulfilled is a new test for models of the evolution of the Kuiper belt.

The third reason for relishing the family's identification is that it gives us some idea about the origin of the spectral differences among Kuiper belt objects. The objects exhibit a variety of colours, from grey to red. A popular model[10] to explain this is that young surfaces, exposed by recent collisions, are grey, whereas older surfaces become gradually reddened by the bombardment of cosmic rays,

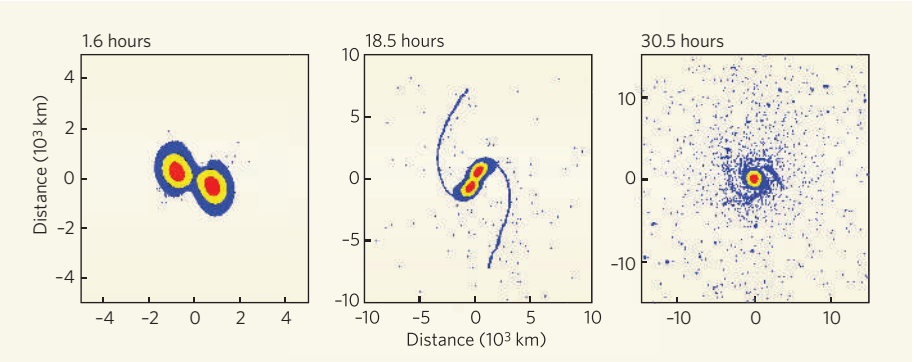


Figure 1 | When worlds collide. Three snapshots from a hydrodynamic simulation[6] show a low-velocity collision between two equally sized objects of 2,000 km diameter. In this case, the collision ejects about 1% of the material into space (blue, ice; yellow, mantle rock; red, iron). This is about 20 times less than the amount required to explain the formation of the 2003 EL_{61} family identified by Brown and colleagues[2]. Further simulations will be needed to explain exactly how this family arose.

a phenomenon known as space weathering. Brown and colleagues' family should be very old, but all its members are grey. This implies that space weathering is not the cause of the colour differences. They could instead reflect compositional differences that depend on the environment in which the objects formed.

For all of the above reasons, Brown and colleagues' identification[2] of this first collisional family in the Kuiper belt would justify its being called a milestone in Kuiper belt science — and, by extension, in our understanding of the outer Solar System's development. ■

Alessandro Morbidelli is at the Laboratoire Cassiopée, Observatoire de la Côte d'Azur, BP 4229, 06304 Nice Cedex 4, France.
e-mail: morby@obs-nice.fr

1. Durda, D. D. *et al. Icarus* **186,** 498–516 (2007).
2. Brown, M. E., Barkume, K. M., Ragozzine, D. & Schaller, E. L. *Nature* **446,** 294–296 (2007).
3. Bendjoya, P. & Zappala, V. in *Asteroids III* (eds Bottke, W. F. *et al.*) 613–618 (Univ. Arizona Press, Tucson, 2002).
4. Michel, P., Benz, W., Tanga, P. & Richardson, D. C. *Science* **294,** 1696–1700 (2001).
5. Rabinowitz, D. L. *et al. Astrophys. J. Lett.* **639,** 1238–1251 (2006).
6. Canup, R. M. & Asphaug, E. *Nature* **412,** 708–712 (2001).
7. Canup, R. M. *Science* **307,** 546–550 (2005).
8. Trujillo, C. A., Jewitt, D. C. & Luu, J. X. *Astron. J.* **122,** 457–473 (2001).
9. Stern, S. A. *Astron. J.* **110,** 856–868 (1995).
10. Jewitt, D. C. & Luu, J. X. *Astron. J.* **122,** 2099–2114 (2001).

CELL BIOLOGY

Lost in mitotic translation

Anthony Wynshaw-Boris

A protein called 14-3-3σ inhibits the cell cycle and may act as a tumour suppressor. It now turns out that it is also involved in regulating protein synthesis from messenger RNA during cell division.

Movement through the cell cycle is regulated by structural modification of proteins after they have been translated from messenger RNA (mRNA). Such changes are known to be important in regulating DNA checkpoint and repair pathways, in triggering cell division (mitosis) and in maintaining the mitotic spindle, which ensures that the chromosomes separate correctly. Less well appreciated is the fact that changes in the step before protein modification — protein translation — are also involved in regulating the progression of the cell cycle[1]. On page 329 of this issue, Wilker *et al.*[2] show that a protein called 14-3-3σ, which is known for its involvement in tumour suppression, is a critical regulator of translation during mitosis, and is essential for normal mitotic progression.

Translation in eukaryotic cells is initiated by two mechanisms. Eukaryotic mRNAs contain a 'cap' consisting of a modified guanosine residue at one end (the 5′ end). Generally, mRNA translation is cap dependent, which means that a protein known as the eukaryotic initiation factor 4B (eIF4B) binds to this cap to initiate mRNA translation. Several viral RNAs do not require a 5′ cap for translation in a eukaryotic cell; translation is instead mediated by so-called internal ribosome entry sites (IRESs) contained in their non-capped mRNAs. This provides a mechanism that favours the translation of viral mRNAs during infection once cap-dependent translation of the host mRNAs has been suppressed.

Cap-independent translation also occurs in the normal eukaryotic cell cycle. During and immediately after the metaphase stage of mitosis, cap-dependent translation is suppressed, and cap-independent translation is stimulated[3]. This is because mRNAs for some cell-cycle proteins contain IRESs that allow their translation during mitosis[4]. The importance of this switch from cap-dependent to cap-independent translation in regulating cell-cycle progression had been unknown. But Wilker *et al.*[2] now provide insights into a potential role of this pathway and of 14-3-3σ in tumorigenesis.

The 14-3-3s are a ubiquitous family of proteins that are evolutionarily conserved from yeast to humans, and they were first identified through their high level of expression in the mammalian brain[5]. Mammals have eight different protein forms of 14-3-3 that are encoded by seven distinct genes (β, ε, γ, η, τ, ζ and σ). More than 200 'client' proteins have been identified that bind to 14-3-3 proteins, and these are involved in a range of cellular processes, including signal transduction, metabolic regulation and trafficking of cellular components. Such processes are involved in cell-cycle responses, programmed cell death, development and cancer.

14-3-3σ is unique among the 14-3-3s. It apparently occurs only in mammals and prefers to form dimers with itself (homodimers) rather than heterodimers with other 14-3-3s. The 14-3-3σ gene was first identified as a gene that responds to the activation of the p53 tumour-suppressor protein after DNA damage. 14-3-3σ was found to bind to essential regulators of the cell cycle, such as cyclin B1 and cdc2, sequestering them in the cytoplasm after DNA damage to prevent them from entering the nucleus where they function[6,7].

In the absence of 14-3-3σ, cells undergo abnormal division after DNA damage, resulting in binucleated and multinucleated cells[7]. Several studies have found that 14-3-3σ is lost or inactivated early in the development of human tumours, including breast and prostate cancers, suggesting that the protein acts as a tumour suppressor[8].

Wilker *et al.*[2] add weight to this idea by showing that, during mitosis, 14-3-3σ binds to several translation initiation factors, including eIF4B. This factor promotes the recruitment of ribosomes — the cell's protein-synthesizing machinery — to mediate cap-dependent translation (Fig. 1a). Furthermore, cells lacking 14-3-3σ do not suppress cap-dependent translation or stimulate cap-independent

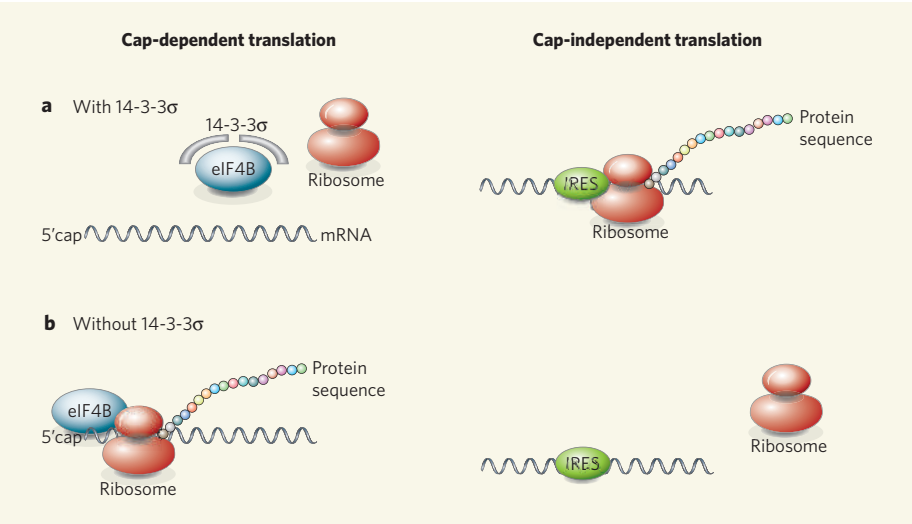


Figure 1 | Role of the 14-3-3σ protein in cap-dependent and cap-independent mRNA translation. During mitosis, 14-3-3σ binds to several proteins required for translation, including eIF4B. This latter protein is required for translating mRNAs containing 5′ caps. **a**, Wilker *et al.*[2] show that, when bound to 14-3-3σ, eIF4B cannot bind to the 5′ cap. Consequently, cap-dependent translation is suppressed and only mRNA sequences with IRES domains, such as the mRNA for cell-cycle regulator p58-PITSLRE, can undergo translation in a cap-independent manner. **b**, Without 14-3-3σ, eIF4B is free to bind to the 5′ cap, initiating cap-dependent translation. Under these circumstances, cap-independent translation is not stimulated.

translation during and immediately after mitosis (Fig. 1b), indicating the protein's crucial role in regulating translation during mitosis. Instead, in cells lacking 14-3-3σ, cell division is impaired, with a loss of the Plk1 protein — a regulator of mitotic progression — at the mid-body region between dividing cells, and an accumulation of binucleated and fused cells.

Wilker and colleagues then searched for proteins that both were translated during mitosis in a cap-independent manner and had been implicated in mitotic events. They identified a proposed target of this defective switch in translation — the cell's own IRES-dependent form of a protein involved in cell-cycle progression (the cyclin-dependent kinase cdk11), called p58-PITSLRE. This is one of a large family of proteins involved in regulating cell-cycle progression. It contains an IRES element in its mRNA-coding region, and its activity is regulated by the cell cycle, thereby permitting translation of p58-PITSLRE immediately before and during mitosis[9]. The evidence for p58-PITSLRE as a target of 14-3-3σ is compelling, but it is unlikely to be the only cap-independent mRNA affected by the defective switch in translation that is regulated by 14-3-3σ. It will be essential to identify other cap-independent mRNAs involved in this response.

Wilker *et al.*[2] also showed that some cellular defects observed after the 'knockdown' of 14-3-3σ can be reversed by using the drug rapamycin to suppress cap-dependent translation and increase cap-independent translation during mitosis. The authors therefore argue that, in the absence of 14-3-3σ, aberrant mitotic translation impairs termination of mitosis and produces binucleated cells; this implies that 14-3-3σ may be involved in tumorigenesis.

The identification of 14-3-3σ as a protein required for the mitotic switch from cap-dependent to cap-independent translation shines a brighter light on mRNA translation as an underappreciated factor in cell-cycle control. These findings also suggest that transition from cap-dependent to cap-independent translation is a potential chemotherapeutic target, as the effects of rapamycin may be partly explained by its ability to alter translation. Thus, there could be potential in exploiting the mitotic regulation of mRNA translation for therapeutic purposes. ■

Anthony Wynshaw-Boris is in the Departments of Pediatrics and Medicine, Center for Human Genetics and Genomics, University of California, San Diego, La Jolla, California 92093-0627, USA.
e-mail: awynshawboris@ucsd.edu

1. Holcik, M. & Sonenberg, N. *Nature Rev. Mol. Cell Biol.* **6,** 318–327 (2005).
2. Wilker, E. W. *et al. Nature* **446,** 329–332 (2007).
3. Pyronnet, S. *et al. Genes Dev.* **15,** 2083–2093 (2001).
4. Pyronnet, S. *et al. Mol. Cell* **5,** 607–616 (2000).
5. Darling, D. L., Yingling, J. & Wynshaw-Boris, A. *Curr. Top. Dev. Biol.* **68,** 281–315 (2005).
6. Hermeking, H. *et al. Mol. Cell* **1,** 3–11 (1997).
7. Chan, T. A. *et al. Nature* **401,** 616–620 (1999).
8. Lodygin, D. & Hermeking, H. *Cell Res.* **15,** 237–246 (2005).
9. Cornelius, S. *et al. Mol. Cell* **5,** 597–605 (2000).

QUANTUM PHYSICS

Total surveillance

Ferdinand Schmidt-Kaler

Trapped by mirrors, a photon can be monitored from birth to death by a stream of passing atoms. The technique could also be used to entangle the quantum states of many atoms — a possible boon for quantum computing.

"You measured this only once?" Such enquiries from professor to student are often heard in the experimental laboratory. But although the necessity of repeat measurements might be self-evident in the classical world, things are not quite so straightforward for quantum measurements. This is because the tiny objects of quantum desire — single photons, for instance — are easily destroyed simply by measuring them. On page 297 of this issue, Serge Haroche and colleagues (Gleyzes *et al.*)[1] observe for the first time a single photon in a sequence of about 100 consecutive measurements, recording it from the instant of its birth to the moment of its death.

What are the necessary ingredients for this experimental masterwork? First, you must trap a photon for long enough to measure it over and over again. The longer the photon is trapped, the better. In Gleyzes and colleagues' experiment, a cavity in the form of a box 2.7 cm long, with walls made from ultra-reflective, superconducting mirrors (Fig. 1), serves to hold a photon for about 0.13 seconds. This is an impressive achievement: in that same time, a free photon would travel about a tenth of the distance to the Moon.

Second, a 'transparent' counter must be used instead of a normal photodetector. A conventional light detector works by absorbing the energy of an incident photon, and in so doing destroys it. But even if the photon is not destroyed during the measurement process, its properties will change dramatically. Gleyzes *et al.* use what is known as a quantum non-demolition technique. This involves a stream of atoms crossing the box in which the photon is trapped. Light is an electromagnetic wave, and the electric field of the photon shifts the energy levels of the atom, but without giving the atom a chance to absorb energy from the field. Once the atoms have emerged from the interaction region, they can be analysed precisely to ascertain whether a photon was in the box or not.

The principle of this measurement process is akin to each atom carrying its own clock, with orbiting electrons acting as the pendulum. The frequency of this pendulum is slowed down by the photon's electric field while the atom is in the box with a photon. This creates a delay in the atom's clock when compared with the clock of an atom that has not interacted with a photon. Measuring this minuscule difference requires the sensitive technique, devised by Norman Ramsey[2], that is widely used for receiving the signal of modern atomic clocks.

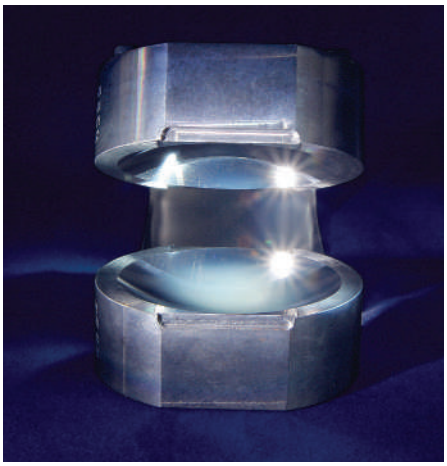


Figure 1 | They do it with mirrors. Gleyzes and colleagues' cavity for trapping photons[1].

We must still remember, however, that according to the strict laws of the quantum world, any measurement of a photon causes an unavoidable disturbance. For a classical electric field, both phase and amplitude may be determined independently with infinite precision. By contrast, the fluctuation in the phase and amplitude of a quantum field — such as that of the trapped photon — are interconnected. These two 'non-commuting' variables obey the Heisenberg uncertainty relation: that is, increased precision in measuring one of them will reduce the precision in measuring the other. In Gleyzes and colleagues' experiments[1], the amplitude of the photon field is determined with certainty, meaning necessarily that the phase of the photon wave is fully mixed up.

Roy Glauber[3] was the first to develop the quantum theory of photon detection in a destructive measurement process. Gleyzes *et al.* illustrate beautifully the detection of a single photon without its destruction. But their experimental scheme is not limited to simple cases such as that of a single photon, and further exciting experimental news is to be expected. With their current set-up, the authors are already in a position to detect quantum superpositions of two distinct, classical states of the photon's electric field. The photon statistics of such notorious 'Schrödinger's cat' states have escaped observation until now.

The successful demonstration[1] of a quantum non-demolition measurement of a single photon has significant implications for the rapidly evolving field of quantum computing[4].

SOCIAL SCIENCE

The urban organism

Visitors to the area around *Nature*'s London offices will be familiar with the scene: unending traffic and noise; the hurly-burly of the Underground; streets, concourses and platforms filled with people intent only on reaching their destination quickly. It's received wisdom that the bigger a city is, the faster life moves; Luis Bettencourt and colleagues supply some empirical evidence to back up that perception (*Proc. Natl Acad. Sci. USA* **104,** 7301–7306; 2007).

The authors begin by examining how different indicators of cities' activity and infrastructure scale with their size. They use various sets of data from the United States, China and Germany, and characterize the scalings as power laws of the form (population)^n. They find that indicators of economic activity — from personal income, to patent registrations, to total electricity consumption — vary with population with values of n in the range 1.1–1.3, regardless of where the data were collected. In other words, cities the world over become more hyperactive the larger they get. Perhaps as a corollary to that excess, the prevalence of crime and sexually transmitted disease grows similarly quickly.

Infrastructure indicators such as the lengths of the road and electricity networks, by contrast, scale to around (population)$^{0.8}$. The larger the metropolis, the less of these things each citizen has at their disposal. Thus it seems that cities fulfill two basic needs of modern human society: they facilitate the exchange of ideas and, by extension, wealth creation; and they achieve economies of scale in the supply of a population's needs.

To look at how these very different dynamics affect city expansion over time, Bettencourt *et al.* construct a general equation that models the cost on resources of sustaining and increasing a population. Unsurprisingly, growth driven by the demands of efficiency, $n < 1$, stagnates after time: economies of scale eventually hit a bottom line.

City growth driven by wealth creation ($n > 1$), on the other hand, rapidly becomes hyperexponential. The only way to avoid collapse as a population outstrips the finite resources available to it is through constant waves of innovation. These effectively re-engineer the initial conditions of growth. But the greater the absolute population, the smaller the relative return on each such investment — so new ideas must come ever faster.

The city dweller looking for a quiet life is thus hit with a double whammy: the bigger the city, the faster life is; but the rate at which life gets faster must itself accelerate to maintain the city as a going concern.

In biological organisms, the authors note, the situation is completely different. Larger organisms have greater economies of scale, and slower-paced lives. Metabolic rates, for example, increase with (body mass)$^{0.75}$. With the city, it seems, mankind has created an organism operating beyond the bounds of what is natural. ■

Richard Webb

addition of a ubiquitin chain to Cdc20 by APC/C and UbcH10 does not necessarily involve protein degradation, but leads to the dissociation of Mad2 and BubR1 from Cdc20. One implication of this model is that APC/C would constantly antagonize its inhibition by the spindle-assembly checkpoint. If so, how could Mad2 and BubR1 ever inhibit the APC/C in cells with an active checkpoint?

A possible answer comes from a study carried out by Elledge's team[1]. In a search for proteins that are required for the activity of the spindle-assembly checkpoint, these authors identified a de-ubiquitinating enzyme known as USP44. Enzymes of this type disassemble ubiquitin chains by cleaving the bonds that connect the ubiquitin residues in the chain. Interestingly, USP44 differs from other known spindle-assembly checkpoint proteins in that it is not required to recruit Mad2 to unattached kinetochores, where Mad2 is believed to form complexes with Cdc20 and BubR1. So how else could USP44 function at the checkpoint? It turns out that, *in vitro*, USP44 can inhibit the ability of UbcH10 to activate checkpoint-inhibited APC/C, leading Elledge and colleagues to propose that USP44 might stabilize Cdc20–Mad2–BubR1 complexes by destroying the ubiquitin chains that APC/C adds to Cdc20 (Fig. 1b). Consistent with this argument, depletion of USP44 prematurely inactivates the spindle-assembly checkpoint in mitotic cells and leads to defects in chromosome segregation.

The model proposing that the stability of Cdc20–Mad2–BubR1 complexes is controlled by a fine balance between ubiquitination, mediated by the APC/C, and de-ubiquitination, catalysed by USP44, makes a number of predictions. Testing these will be an essential goal for the future. For example, could a mutant of Cdc20 be created that couldn't be ubiquitinated but would otherwise be functional? If so, such a mutant would be predicted to assemble into unusually stable checkpoint complexes from which Mad2 and BubR1 could not easily dissociate.

These studies[1,2] also raise a number of other questions. Is de-ubiquitinating Cdc20 the main role of USP44 in maintaining the spindle-assembly checkpoint, or does it also antagonize APC/C more directly by disassembling ubiquitin chains on its other protein substrates such as securin and cyclin B? How does ubiquitination dissociate Mad2 and BubR1 from Cdc20 — by inducing conformational changes in these proteins, or by recruiting enzymes (such as the p97/Cdc48–ATPase) that would catalyse the dissociation process? Finally, how is the balance between de-ubiquitination and ubiquitination reactions tipped once all chromosomes have become attached to both poles of the mitotic spindle? Elledge *et al.* found that USP44 itself is degraded at the end of cell division. Could this be the primary switch for checkpoint inactivation, or is it merely a consequence of APC/C activation once the checkpoint has been silenced? Answering these questions will keep researchers busy for some time to come. ■

Jan-Michael Peters is at the Research Institute of Molecular Pathology (IMP), Dr. Bohr-Gasse 7, A-1030 Vienna, Austria.
e-mail: peters@imp.univie.ac.at

1. Stegmeier, F. *et al. Nature* **446,** 876–880 (2007).
2. Reddy, S. K., Rape, M., Marganski, W. A. & Kirschner, M. W. *Nature* **446,** 921–925 (2007).
3. Rieder, C. L., Schultz, A., Cole, R. & Sluder, G. *J. Cell Biol.* **127,** 1301–1310 (1994).
4. Nasmyth, K. *Cell* **120,** 739–746 (2005).
5. Li, Y., Gorbea, C., Mahaffey, D., Rechsteiner, M. & Benezra, R. *Proc. Natl Acad. Sci. USA* **94,** 12431–12436 (1997).
6. Hwang, L. H. *et al. Science* **279,** 1041–1044 (1998).
7. Kim, S. H., Lin, D. P., Matsumoto, S., Kitazono, A. & Matsumoto, T. *Science* **279,** 1045–1047 (1998).
8. Peters, J.-M. *Nature Rev. Mol. Cell Biol.* **7,** 644–656 (2006).
9. Sudakin, V., Chan, G. K. & Yen, T. J. *J. Cell Biol.* **154,** 925–936 (2001).
10. Luo, X., Tang, Z., Rizo, J. & Yu, H. *Mol. Cell* **9,** 59–71 (2002).
11. Sironi, L. *et al. EMBO J.* **21,** 2496–2506 (2002).

Correction
In the News & Views article "Cell biology: Lost in mitotic translation" by Anthony Wynshaw-Boris (*Nature* **446,** 274–275; 2007), statements in the text and the caption to Figure 1, and part **b** of Figure 1, imply that eIF4B binds directly to the 5′ cap of mRNA. Rather, eIF4B facilitates the ATP-dependent helicase activity of eIF4A to promote the ribosome recruitment necessary for cap-dependent translation.

Photons[5], atoms[6], ions[7–9] and also a menagerie of solid-state two-level systems[10–12] have been used to carry elementary units of quantum information ('qubits'), which live in superpositions of two quantum states, 0 and 1. To translate Gleyzes and colleagues' experiment into the language of information processing, they have demonstrated that a stream of atomic qubits can be fully controlled by the qubit state of a trapped photon. In the next experimental steps, we can hope to see the electromagnetic field being prepared in a quantum superposition of one photon and no photon. In this case, the atoms in the stream leaving the box will all be in that same superposition state. Moreover, the properties of all the atoms will be 'entangled' such that it no longer makes any sense to consider them individually, even if they are spatially well separated.

As Nobel-prizewinning physicist Richard Feynman said, "We do not understand quantum mechanics" — but, as Gleyzes and colleagues' achievement[1] shows, we do know increasingly well how to handle it in our experiments. ■

Ferdinand Schmidt-Kaler is at the Institut für Quanteninformationsverarbeitung, Universität Ulm, Albert-Einstein-Allee 11, D-89069 Ulm, Germany.
e-mail: ferdinand.schmidt-kaler@uni-ulm.de

1. Gleyzes, S. *et al. Nature* **446,** 297–300 (2007).
2. http://nobelprize.org/nobel_prizes/physics/laureates/1989/index.html
3. http://nobelprize.org/nobel_prizes/physics/laureates/2005/index.html
4. Bruß, D. & Leuchs, G. (eds) *Lectures on Quantum Information* (Wiley-VCH, Berlin, 2006).
5. Nogues, G. *et al. Nature* **400,** 239–242 (1999).
6. Mandel, O. *et al. Nature* **425,** 937–940 (2003).
7. Häffner, H. *et al. Nature* **438,** 643–646 (2005).
8. Leibfried, D. *et al. Nature* **438,** 639–642 (2005).
9. Schmidt-Kaler, F. *et al. Nature* **422,** 408–411 (2003).
10. Kane, B. E. *Nature* **393,** 133–137 (1998).
11. Nakamura, Y., Pashkin, Yu. A. & Tsai, J. S. *Nature* **398,** 786–788 (1999).
12. Childress, L. *et al. Science* **314,** 281–285 (2006).

THEORETICAL CHEMISTRY

The six-bond bound

Gernot Frenking and Ralf Tonner

What is the maximum number of covalent chemical bonds that two atoms can share? Six, according to the latest theoretical study — at least where just two atoms of the same element are concerned.

Every atom can make a small, integer number of covalent chemical bonds with neighbouring atoms. This number is usually not more than eight, although a precise statement is dangerous — in rare cases, atoms can have more than eight neighbours, but whether they are covalently bound is open to debate. By the nineteenth century, chemists had realized[1] that more than one bond may connect the same two atoms: most famously, adjacent carbon atoms can be bound twice or even three times. For many years, three seemed to be the limit. Then, in 1965, a salt compound was isolated with a fourfold rhenium–rhenium bond in its anion[2]. Five bonds followed in 2005, when the synthesis[3] of a complex organometallic compound was interpreted in terms of a quintuple bond between two chromium atoms[4,5]. In a thorough theoretical study in *Angewandte Chemie International Edition*, Roos, Borin and Gagliardi[6] use quantum chemistry to tackle the obvious question that arises — what is the maximum number of bonds that can bind two atoms?

Theoretical models of covalent bonding go back to 1916, when Gilbert Lewis[7] introduced the idea of electron pairing between neighbouring atoms. Remarkably, this simple model is still the most widespread conceptual description of the covalent bond[8]. The first quantum-chemical study[9,10], however, struggled to reconcile the paired-electron idea with the delocalized nature of electron wavefunctions, which are smeared over large areas of an atom in 'orbitals' (Box 1). Since then, the paired-electron bond has become chemistry's unicorn — although everyone seems to know what one looks like, no one has actually ever seen one[11].

Roos *et al.*[6] simplify things by focusing mainly on molecules made up of two atoms of the same elements, in which there are no further bonds to complicate the analysis. They considered all possible bonding and antibonding combinations of the two atoms' electron orbitals using 'multi-reference' wavefunctions. Here, the simple picture of molecular orbitals that are either empty or occupied by two electrons, one from each atom, does not apply. Instead, the authors calculated 'natural' molecular orbitals that retain the simple model, but have non-integer occupation numbers. By summing up the occupation numbers of all these orbitals (antibonding orbitals being negative in the sum), they could arrive at an overall, non-integer 'effective bond order' (EBO). To tie this in with the qualitative paired-electron bonding picture, the authors suggest that the true 'number' of bonds is the next integer larger than the molecules' EBO.

To find the molecules with the highest EBO, Roos *et al.* investigated the transition-metal dimers of chromium, molybdenum and tungsten, respectively Cr_2, Mo_2 and W_2. The atoms of these transition metals have six outer, or 'valence' orbitals, all of which are available for bonding (Fig. 1). (The availability of an orbital for bonding cannot be assumed: in dimers of atoms such as carbon, for instance, the antibonding combination of one pair of orbitals may be lower in energy than the bonding combination of another, effectively allowing a maximum of just three bonds to form.)

For Cr_2, however, the authors came up with an EBO of 3.5, which equates, following their method, to four paired-electron bonds, rather than the expected six for simple dimers. They explain this discrepancy with an imbalance between the optimal bond length for the outermost 4*s* and 3*d* orbital contributions to the Cr–Cr bond (Box 1). At the equilibrium bonding distance, 1.66 Å, the 4*s* component is already

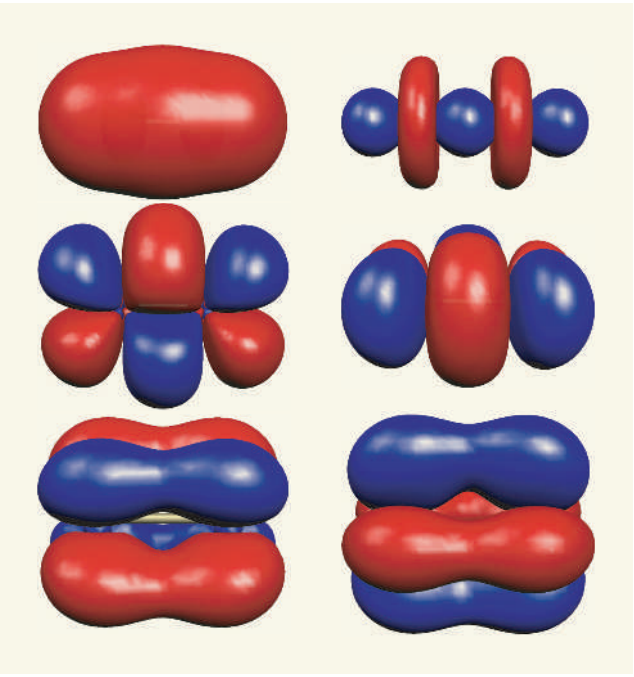


Figure 1 | Six bonds that bind. A depiction of the molecular orbitals of each bond in the sextuple bond between two transition-metal atoms, as investigated by Roos and colleagues[6]. Two bonds (σ bonds; top) each have a single component oriented along the bonding axis between the two atoms, which is horizontal here; two bonds (middle) have double components above and below, or to the sides of, the bonding axis (π bonds); two bonds (bottom) have four components around the axis (δ bonds).

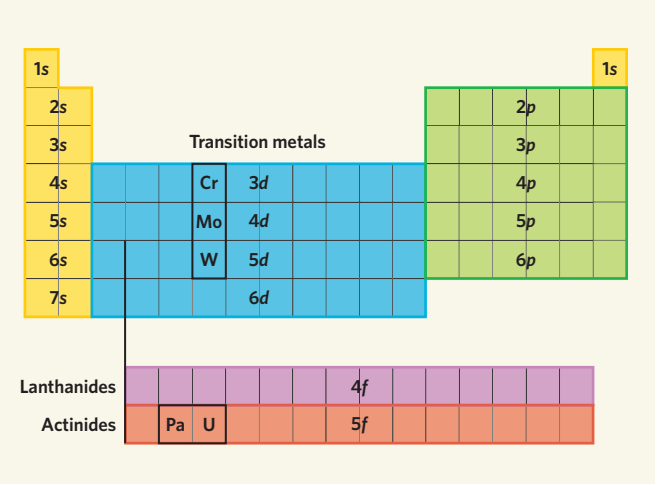
Box 1 | Orbitals and electron counting rules

According to the view prevalent when Gilbert Lewis formulated his picture of covalent bonding[7], electrons circle the atomic nucleus on prescribed orbits. The advent of quantum mechanics, with its probabilistic wavefunctions that smear an electron out over space, modified those classical orbits into quantum 'orbitals'. Each of these describes the spatial distribution of an electron in a particular allowed energy state of the atom.

Two quantum numbers, n and l, are generally sufficient to define an atomic orbital. (A third, m, distinguishes between similar orbitals oriented differently, and a fourth, s, defines the spin of the electron in the orbital.) The principal quantum number, n, depends roughly on the orbital's radial size, and is often thought of as indicating a 'shell' of electrons around the nucleus. The azimuthal quantum number, l, specifies the angular momentum and so the shape of the orbital. Each shell n contains 'subshells' with values of l from 0 to $n-1$. These subshells are represented by letters, the lowest of them reflecting their appearance in early spectroscopic investigations of the atom: *s* (sharp), *p* (principal), *d* (diffuse) and *f* (fundamental).

Taking into account the effects of magnetic field and electron spin, each subshell nl can contain $2(2l+1)$ electrons. As the atomic number increases, electrons enter an orbital in order of ascending energy (the superscript numbers indicate the number of electrons the orbital can hold): $1s^2$, $2s^2$, $2p^6$, $3s^2$, $3p^6$, $4s^2$, $3d^{10}$, $4p^6$, $5s^2$, $4d^{10}$, $5p^6$, $6s^2$, $4f^{14}$, $5d^{10}$, $6p^6$, $7s^2$, $5f^{14}$, $6d^{10}$. That is sufficient to fill the conventional periodic table (pictured), which is structured such that for the elements in the main group the most recently added (outermost) electrons are in the *s* and *p* orbitals. The *d* orbitals make their appearance with transition metals such as chromium (Cr), molybdenum (Mo) and tungsten (W), as investigated by Roos *et al.*[6]. The *f* orbitals appear only with the lanthanides and with actinides such as protactinium (Pa) and uranium (U)[6].

The 8/18 electron counting rules that explain the stability of many molecules are a further consequence of this orbital order.



When elements combine to form compounds, their atomic orbitals combine to form molecular orbitals. Compounds of elements from the main group of the periodic table are particularly stable when they have eight electrons in their outermost shell, corresponding to fully occupied $(n)s^2(n)p^6$ orbitals. This octet rule was another landmark contribution by Lewis[7,8]. A similar 18-electron rule that applies for transition-metal compounds with full $(n)s^2(n)p^6(n-1)d^{10}$ orbitals was formulated by Irving Langmuir[16] in 1921.

Dognon *et al.*[17] now tentatively extend this rule to 32 electrons. They investigated icosahedral compounds in which a metal atom is trapped inside an anionic structure of 12 lead atoms, Pb_{12}^{2-}. When the trapped metal atom was a plutonium ion, Pu^{2+} — with an overall completely filled $(n)s^2(n)p^6(n-1)d^{10}(n-1)f^{14}$ 32-electron system — a perfect match could be made. Time will tell how general this concept is. **G.F. & R.T.**

repulsive, whereas the 3*d* bonding has not yet reached its maximum strength. Significant repulsion between the outermost core electrons of the Cr atoms also occurs at this distance, weakening the bond still further. The dissociation energy needed to break the Cr–Cr bond, 1.65 electronvolts, is therefore low. But as Roos *et al.*[6] point out, this dissociation energy depends on many factors besides the bond order. These include the interplay between attractive and repulsive electrostatic interactions between electrons and nuclei, and the quantum-mechanical Pauli repulsion between electrons that have the same spin[12].

The authors' calculations for Mo_2 and W_2, which are heavier homologues of Cr_2, give natural bond orders of 5.2, suggesting the presence of sextuple bonds[5]. In these dimers, the equilibrium bonding distance — 1.95 Å in Mo_2, 2.01 Å in W_2 — is greater than in Cr_2. The outermost orbitals (5*s* and 4*d* in Mo_2, 6*s* and 5*d* in W_2) are also better balanced and have a much larger radial extent[13]. The resulting more effective overlap of orbitals between the two atoms yields not only a higher bond order, but also much stronger bonds, with dissociation energies of 4.4 and 5.4 eV, respectively. This trend of bond strength increasing with higher atomic weight is quite general among the transition metals, where relativistic effects have an important role[14]; the trend for the elements in the main group of the periodic table is usually the reverse.

Roos *et al.*[6] also calculated the EBO of certain transition-metal compounds. For the archetypal $Re_2Cl_8^{2-}$ anion — that in which a quadruple bond was first identified[2] — their EBO of 3.2 supports the standard interpretation[15]. But analysis of an organometallic model compound structurally similar to that in which the quintuple Cr–Cr bond was thought to have been spotted[3] yields an EBO of just 3.5.

Is it possible for a covalent bond to have an order greater than six? In principle, this should be possible for bonds between atoms such as the lanthanides and actinides. These each have seven atomic *f* orbitals that can yield seven *f*–*f* combinations. But the 4*f* orbitals in lanthanides are known to be much more contracted than the 6*s* and 5*d* orbitals that are the next lowest and highest in energy, respectively, meaning that they do not participate in chemical bonding. In actinides, Roos *et al.* do find a weak participation of the equivalent 5*f* orbitals in diatomic species M_2, but the calculated EBOs — at most, 4.5 for a dimer of the rather exotic element protactinium (Pa), followed by 4.2 for uranium (U) — are far from those of a septuple bond.

The general conclusion thus seems to be that — at least as far as covalent bonds between equal atoms are concerned — six is the limit. As Roos and colleagues admit, however, a measure of bond order between unequal atoms requires a more developed model that includes electrostatic interactions. Whether in these circumstances the upper limit remains bound at six is still open to question. ■

Gernot Frenking and Ralf Tonner are in the Fachbereich Chemie, Philipps-Universität Marburg, Hans-Meerwein-Straße, D-35043 Marburg, Germany.
e-mail: frenking@chemie.uni-marburg.de

1. Mierzecki, R. *The Historical Development of Chemical Concepts* (Kluwer, Dordrecht, 1990).
2. Cotton, F. A. & Harris, C. B. *Inorg. Chem.* **4,** 330–333 (1965).
3. Nguyen, T. *et al. Science* **310,** 844–847 (2005).
4. Brynda, M., Gagliardi, L., Widmark, P.-O., Power, P. P. & Roos, B. O. *Angew. Chem. Int. Edn* **45,** 3804–3807 (2006).
5. Frenking, G. *Science* **310,** 796–797 (2005).
6. Roos, B. O., Borin, A. C. & Gagliardi, L. *Angew. Chem. Int. Edn* **46,** 1469–1472 (2007).
7. Lewis, G. N. *J. Am. Chem. Soc.* **38,** 762–785 (1916).
8. Frenking, G. & Shaik, S. (eds) *90 Years of Chemical Bonding; J. Comput. Chem.* **28** (2007).
9. Heitler, W. & London, F. *Z. Phys.* **44,** 455–472 (1927).
10. Frenking, G. *Theor. Chem. Acc.* **103,** 177–179 (2000).
11. Frenking, G. & Krapp, A. *J. Comput. Chem.* **28,** 15–24 (2007).
12. Krapp, A., Bickelhaupt, F. M. & Frenking, G. *Chem. Eur. J.* **12,** 9196–9216 (2006).
13. Kutzelnigg, W. *Angew. Chem. Int. Edn Engl.* **23,** 272–295 (1984).
14. Pyykkö, P. *Chem. Rev.* **88,** 563–594 (1988).
15. Cotton, F. A., Murillo, C. A. & Walton, R. A. *Multiple Bonds Between Metal Atoms* (Springer, Heidelberg, 2005).
16. Langmuir, I. *Science* **54,** 59–67 (1921).
17. Dognon, J.-P., Clavaguéra, C. & Pyykkö, P. *Angew. Chem. Int. Edn* **46,** 1427–1430 (2007).

REVIEWS

Hybrid speciation

James Mallet[1]

Botanists have long believed that hybrid speciation is important, especially after chromosomal doubling (allopolyploidy). Until recently, hybridization was not thought to play a very constructive part in animal evolution. Now, new genetic evidence suggests that hybrid speciation, even without polyploidy, is more common in plants and also animals than we thought.

Linnaeus stated in *Systema Naturae* that species have remained unchanged since the dawn of time, but he later experimented with hybrids and convinced himself that hybridization provided a means of species modification. One hundred and eighty years later, Lotsy[1] still argued that species were invariant genetic types, and that novel lineages could evolve only by means of hybridization. These peculiar ideas were overturned when the concept emerged of species as reproductively isolated populations[2–4]. In zoology, this concept discouraged the view that hybridization and gene flow (introgression) between species could be important evolutionary forces[2,5,6], even while botanists continued to argue for their significance[4,7,8]. Today, armed with new and abundant molecular marker data, biologists increasingly find new examples where hybridization seems to facilitate speciation and adaptive radiation in animals, as well as plants[8–12].

What is hybrid speciation?

'Hybrid speciation' implies that hybridization has had a principal role in the origin of a new species. The definition applies cleanly to hybrid species that have doubled their chromosome number (allopolyploidy): derived species initially contain exactly one genome from each parent, a 50% contribution from each, although, in older polyploids, recombination and gene conversion may eventually lead to unequal contributions. Furthermore, allopolyploids are largely reproductively isolated by ploidy. Recombinational hybrid speciation, in which the genome remains diploid (homoploid hybrid speciation), is harder to define. The fraction from each parent will rarely be 50% if backcrossing is involved. Homoploid hybrid species may be only weakly reproductively isolated, and are hard to distinguish from species that gain alleles by hybridization and introgression, or from persistent ancestral polymorphisms. Although hybrid speciation is sometimes inferred if any marker alleles originate from different parents, I here restrict the term to cases where hybrid allelic combinations contribute to the spread and maintenance of stabilized hybrid lineages generally recognized as species.

This raises the question of what exactly we mean by 'species'. Hybrid speciation is only possible if reproductive isolation is weak; if hybrids are intermediate, hybrid species will be even more weakly isolated. In practice, we must recognize species as multi-locus 'genotypic clusters' (Box 1)[6,13]. A hybrid species will then be a third cluster of genotypes, a hybrid form that has become stabilized and remains distinct when in contact with either parent.

Hybridization can also influence speciation by means of 'reinforcement', where mating barriers evolve owing to selection against unfit hybrids[6,14,15]. Although hybridization contributes to speciation, I do not consider reinforcement to be hybrid speciation, because a third species does not form. A related and highly relevant phenomenon is 'hybridogenesis'. The diploid or triploid edible frog *Rana esculenta* is a well known example: it is heterozygous for complete *Rana lessonae* and *Rana ridibunda* genomes[16]. Here, I exclude hybridogenetic species because they do not breed true.

Theory and background of hybrid speciation

Hybridization may be "the grossest blunder in sexual preference which we can conceive of an animal making"[17], but it is nonetheless a regular event. The fraction of species that hybridize is variable, but on average around 10% of animal and 25% of plant species are known to hybridize with at least one other species[18]. Hybridization is especially prevalent in rapidly radiating groups: 75% of British ducks (Anatidae)[18], for example. Recent, closely related species are most likely to hybridize, although hybridization and introgression

Box 1 | Species as genotypic clusters versus reproductively isolated populations

Species can be defined as distinguishable groups of genotypes that remain distinct in the face of potential or actual hybridization and gene flow[6,13]. This is similar to Darwin's usage of species to divide biodiversity by means of gaps or troughs in the distributions of phenotypes and genotypes. A very likely reason why a pair of genotypic clusters in contact might remain distinct is, of course, 'reproductive isolation', but this becomes a means of achieving speciation and species maintenance rather than a definition of the species state itself. It might seem that autopolyploids (species resulting from chromosome doubling within a single parent species) cannot be recognized as genotypic clusters because they have initial gene frequencies like their diploid parents. Autopolyploids are, however, genetically distinct in heritable traits such as chromosome number and ploidy at individual loci. They can be regarded as distinct species provided that euploids (for example, diploids and tetraploids) form clusters that are more abundant than intermediates formed by hybridization between them (for example, triploids and aneuploids).

Such species have no guarantee of permanence. Two genotypic clusters might be stable for a long time, yet when ecological circumstances change, gene flow may exceed some threshold, eventually resulting in a single genotypic cluster that absorbs both species[13]. There are several examples of species fusion in the literature, for example in Darwin's finches and cichlid fish[44,45]. 'Despeciation' itself could be classified as a form of hybrid speciation, as a new species has resulted from the fusion of two old species. I exclude despeciation here because, in my definition, hybrid species should remain distinct when in contact with either parent. Many have argued that permanent divergence is an important criterion of species[2]. However, dropping this a priori requirement seems reasonable to avoid the need to predict an often unpredictable future for distinct, existing taxa[6], and allows for extinction of species via genomic swamping, which seems as valid and potentially important as other forms of extinction, especially in human-altered environments[8].

[1]Galton Laboratory, Department of Biology, University College London, 4 Stephenson Way, London NW1 2HE, UK.

may often persist for millions of years after initial divergence[18]. Hybridization is thus a normal feature of species biology, if at a rather extreme end of the natural spectrum of sexuality[5]; it is not merely an unnatural "breakdown of isolating mechanisms"[2]. At the population level, interspecific hybrids are, of course, unusual, forming <0.1% of individuals in a typical population[2,18]; they are also 'hopeful monsters', with hefty differences from each parent, no adaptive history to any ecological niche, and little apparent scope for survival (Fig. 1). Furthermore, hybrids are often sterile or inviable owing to divergent evolution in each species[2,6]. Even if a healthy hybrid is formed, it normally suffers 'minority cytotype disadvantage' because it encounters few mates of its own type, and backcrosses to the more abundant parent species will often be unfit. For example, a rare tetraploid hybrid will produce unfit triploid progeny with diploid parents[19].

Yet hybrid species exist. What advantages could outweigh the catalogue of difficulties? This innocent-sounding question plunges to the heart of controversies about adaptive evolution. Is saltational evolution possible? Are maladaptive intermediates and genetic drift involved? Common sense and prevailing opinion suggests that evolution normally occurs by small adjustments rather than saltation, and rarely involves maladaption[6]. It is therefore extraordinary that hybrid speciation can disobey both rules. Hybridization (or hybridogenesis) can act as a multi-locus 'macro-mutation' that reaches out over large phenotypic distances[5] to colonize unoccupied ecological niches or adaptive peaks (Fig. 1). Furthermore, random drift in small, localized hybrid populations provides a parsimonious solution to maladaptation, to enable local establishment, stabilization and ultimate spread[20].

Two principal types of hybrid speciation are treated here: allopolyploidy and homoploid hybrid speciation.

Hybrid speciation through allopolyploidy

Polyploidy is a well-established speciation mode in plants, although many aspects of polyploid evolution are only today being revealed[3,19,21,22]. Speciation can be via autopolyploidy (duplication of chromosomes within a species) or allopolyploidy (duplication of chromosomes in hybrids between species), although the boundary is blurred because of the 'fuzzy' nature of species. Polyploid species are reproductively isolated from their parents because when polyploids mate with diploids, progeny with odd-numbered ploidies, such as triploids, are produced. These offspring may be viable but typically produce sterile gametes with unbalanced chromosomal complements (aneuploidy)[3,4,22]. Polyploidy is thus a simple saltational means of achieving speciation[4]. The process may be repeated many times, leading to lineages with >80-fold ploidy in some vascular plants; 40–70% of all plant species are polyploids[3,21].

Allopolyploid speciation can result from somatic chromosome doubling in a diploid hybrid, followed by selfing to produce a tetraploid. This was the route taken by *Primula kewensis*, the allopolyploid that arose spontaneously in 1909 among cultivated diploid hybrids of *Primula verticillata* and *Primula floribunda*[22]. However, there are other possibilities, such as fusion of two unreduced gametes

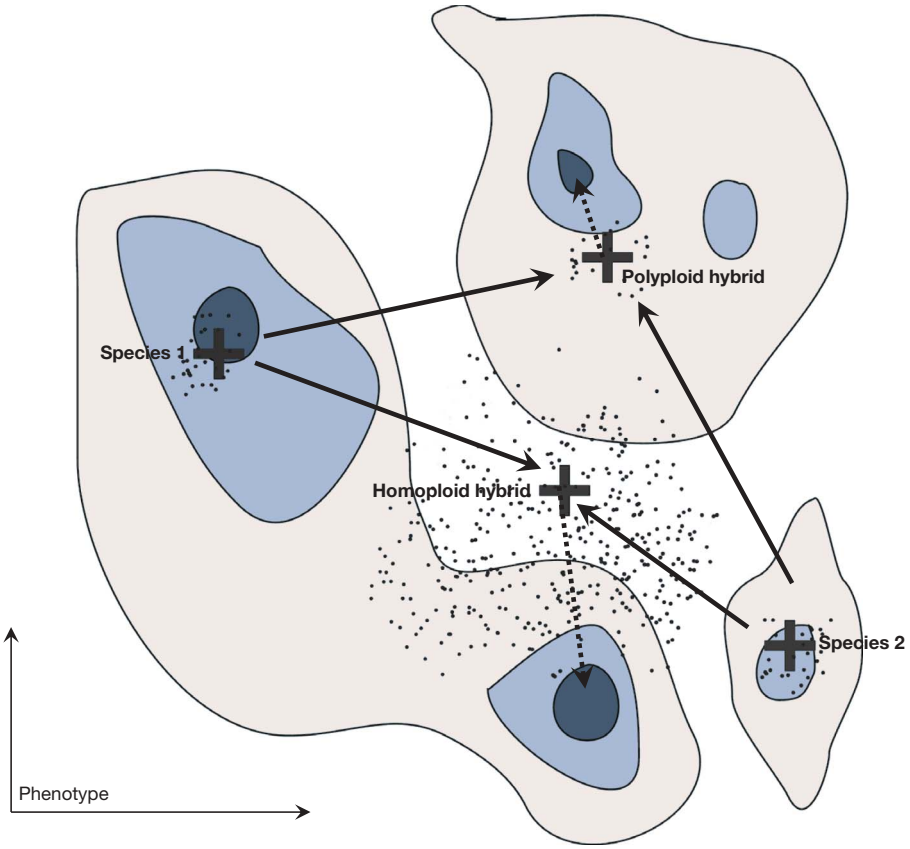


Figure 1 | Hybridization and the adaptive landscape. The hyperspace of possible phenotypes and genotypes can be represented as an adaptive landscape[20]. Fitness optima ('adaptive peaks') are coloured blue. Adaptive landscapes are not rigid, but are readily distorted by environmental or biotic changes, including evolutionary change. Mean phenotypes of species and their hybrids are shown as crosses, and offspring distributions as dots. Species 1 and 2 are adapted to different fitness optima. Natural selection acts mainly within each species, so hybrids are 'hopeful monsters', far from phenotypic optima (solid arrows). It is therefore hard to imagine how hybrids often attain new optima unless unoccupied adaptive peaks are abundant. Polyploid hybrids can have a variety of advantages over their parents, including heterozygote advantage, extreme phenotypic traits and reproductive isolation. Genetic variation in their offspring will initially be similar to that of non-hybrid parents if recombination between parental genomes is rare[4,22]; such hybrids will not spread unless already near an optimum. Homoploid hybrids have fewer initial advantages, but their progeny can have extremely high genetic variances via recombination, including phenotypes more extreme than either parent—transgressive variation (not shown here). This burst of variation can help homoploids attain new adaptive peaks (dotted arrow) far from parental optima[30,32].

after failure of reduction divisions in meiosis. A third route is the 'triploid bridge', in which rare, unreduced (diploid) gametes fuse with normal haploid gametes to form triploids. Triploids are normally sterile, but can contribute to tetraploid formation by themselves producing occasional, unreduced triploid gametes that can backcross with a normal haploid gamete to form tetraploid progeny[19,22]. This was the route used to engineer the first, and maybe the only, synthetic (that is, in the laboratory), self-sustaining bisexual animal polyploid strain, a hybrid between silk moths (*Bombyx mori* and *Bombyx mandarina*)[23].

After polyploid hybrids arise, they still face major hurdles. Diploid and triploid hybrids are strongly disfavoured because their aneuploid gametes are almost always sterile. Even when even-numbered allopolyploidy is achieved, chromosome pairing is rarely perfect[22]. Furthermore, assuming new polyploids are rare, they will mate mostly with incompatible parentals, leading to minority cytotype disadvantage[19]. These problems almost certainly explain why bisexual polyploid speciation is more common in plants than animals: (1) plants usually have indeterminate growth, and somatic chromosome doubling can lead to germline polyploidy; (2) plants are also often perennial or temporarily clonal, allowing multigenerational persistence of hybrid cell lines within which polyploid mutations can occur; (3) plants are more often hermaphrodites, allowing selfing as a means of sexual reproduction of rare polyploids, once formed; and (4) gene flow is weaker in plants than in animals, and local populations with unusual ploidy (whether by local drift or selection) can form more readily to overcome minority cytotype disadvantage. As expected, polyploidy is strongly associated with asexual reproduction, selfing and perenniality in plants, as well as in animals[11,21,23]. In clonal polyploid animals, reversion to out-crossing is rare, whereas in plants, with frequent alternation between clonal and sexual phases, bisexual polyploid species are common and themselves often give rise to further species. Thus, animal allopolyploids such as stick insects (*Bacillus*) and freshwater snails (*Bulinus truncatus*) are often, although not always, parthenogenetic or selfers[21]. Muller's theory[24] that sex chromosomes in animals prevent sexual polyploidy owing to sex:autosome gene dosage is no longer given much credence[6,21,23,25].

How common is polyploid speciation? Otto and Whitton[21] provided new insights from the over-representation of even-numbered chromosome counts. Recent polyploidy explains ~2–4% of speciation events in flowering plants and ~7% of speciation events in ferns[21], and these are probably underestimates[6] (40–70% of plant species overall are polyploid, but this includes the effects of much non-polyploid speciation within already polyploid lineages[3,21]). In animals, there is no bias towards even-numbered chromosomal counts, suggesting that animal polyploid speciation is very rare compared with other speciation modes[21].

Traditional dogma has it that allopolyploids arise more readily than autopolyploids because the latter are more prone to chromosome pairing problems in meiosis[4]; however, this view is no longer generally accepted. Newly arisen autopolyploids have levels of infertility and aneuploid gametes comparable to those of allopolyploids[22]. Furthermore, many autopolyploids probably lie unrecognized by taxonomy within diploid progenitor species[6]. Yet these discoveries give little insight into a more important question: what fraction of polyploids that spread successfully are allopolyploids? Autopolyploids may often be doomed to extinction, perhaps through competition with similar diploid relatives[6]. Opinions differ, but it probably remains true that allopolyploids are more successful than autopolyploids[6,26]; certainly allopolyploids are a sizeable fraction of well-studied crop cases, such as wheat, cotton and tobacco[7,26]. There are almost no surveys of entire floras, although a small-scale survey in the United States revealed that 79–96% of 28 polyploid species were allopolyploids[4]. Recently, the Arctic flora was surveyed, in which about 50% of the often clonal or selfing species are polyploids[27]. In Svalbard (Spitsbergen), 78% of the 161 species are polyploid, with the average level of ploidy approximately hexaploid. Every one of the 47 polyploid species studied genetically shows fixed marker heterozygosity, implying 100% allopolyploidy[27]. The Arctic is, of course, an extreme environment, but this remains the most comprehensive survey so far. If Svalbard is typical, most successful polyploids are also hybrid species.

After formation, novel allopolyploids face the usual 'hopeful monster' difficulties (Fig. 1). It helps if they can exploit a new ecological niche that is both vacant and also spatially separated to ameliorate minority cytotype disadvantage. For example, recent allopolyploid hybrids between introduced and native plants have successfully spread from sites of origin (for example, *Senecio cambrensis* in Wales[28] and *Spartina anglica* in England[29]). These invasive allopolyploids were able to exploit vacant ecological roles with relatively little evolutionary change (Fig. 1).

Stochastic drift may also be necessary to overcome minority cytotype and other disadvantages. A few polyploids, usually from the same hybridization event, must accumulate locally for the process to take off, probably involving chance or an unusual local selective regime. Stochastic effects are evident in nature. An independently derived Scottish population of *S. cambrensis* became extinct in Edinburgh some 20 yr after being discovered[28]. Of two origins, only the Welsh population now survives. Other allopolyploid hybrids can arise repeatedly from the same parents[26], but many widespread polyploids (for example, *S. anglica*) probably originated only once or a few times[21,28,29], even though parent species are in broad contact, again showing the importance of chance in the origins of hybrid species.

Recombinational and homoploid hybrid speciation

Homoploid hybrid speciation or recombinational speciation is well-known in flowering plants[4,7,30]. Speciation takes place in sympatry (by definition, as hybridization requires gene flow). Hybrids must then overcome chromosome and gene incompatibilities, while lacking reproductive isolation via polyploidy. For these reasons, the process is often considered unlikely[5,6,31].

However, hybridization can boost genetic variance[30,32], allowing colonization of unexploited niches (Fig. 1). Suppose + and − alleles at genes affecting a quantitative trait differ between species, so that each has fixed differences (+++−− and −−−++, say). Recombination can then liberate 'transgressive' quantitative variation[32], often more extreme than either parent (for example, −−−−− and +++++). Most early recombinants will be unfit, but extreme hybrids can colonize niches unavailable to parents. If ecological opportunities are partially separated from the parental habitat, if like hybrids tend to associate (for example, by means of seasonality or drift in small populations), or if selfing or inbreeding is common, gene flow between hybrids and parents will be reduced and hybrid speciation becomes more likely[7,32]. Successful hybrid species might also displace one or more parent species ecologically[30], and obliterate evidence of their own hybrid origin.

In plants, about 20 well-established homoploid hybrid species are known[31,33], but they are hard to detect and may be more prevalent. The best documented are the desert sunflowers *Helianthus anomalus*, *Helianthus deserticola* and *Helianthus paradoxus*, which all derive from hybrids between mesic-adapted *Helianthus annuus* and *Helianthus petiolaris*[30,33]. Selfing is rare and provides little assistance to establishment, but the three hybrid species survive drought better than their parents, suggesting recruitment of hybrid transgressive variation. Synthetic hybrid populations are readily recreated with karyotypic combinations like those in wild hybrid species, because selection repeatedly favours similar combinations of compatible chromosomal rearrangements. In addition, the wild contribution from each parent of extreme adaptive traits for morphology, physiology and life history of the hybrids (for example, small leaf size, seed dormancy, or tolerance of drought and salt) matches experimental predictions[32,33]. In *Helianthus*, recombinant genotypes and spatial separation have enabled the hybrids to flourish where their parents are absent.

Although bisexual polyploids are often barred in animals, there is no reason why homoploid hybrid species would be rarer in animals

than in plants. The number of cases in animals is growing rapidly[10,33]. A recent example is the invasive sculpin, a hybrid fish derived from the *Cottus gobio* group from the Scheldt River (compare *Cottus perifretum*) and upper tributaries of the Rhine (compare *Cottus rhenanum*). Sculpins are normally restricted to clear, well-oxygenated cold waters in upper river tributaries across Europe. The Rhine and Scheldt rivers became connected as a result of earlier canal building, but invasive sculpins appeared in the warmer and muddier lower Rhine only in the past fifteen years. Morphologically, the invasive sculpin is intermediate, and its mitochondrial DNA, as well as nuclear single nucleotide polymorphisms and microsatellites, are characteristic of both Scheldt and Rhine forms[34]. The hybrid form meets upper Rhine sculpins in narrow hybrid zones, but remains distinct despite gene flow, suggesting that it is adapted mainly to the lower Rhine. Recent evolution and spread of the invasive sculpin, as well as intermediacy, provides convincing evidence of adaptive hybrid origin. A more ancient example is the cyprinid fish *Gila seminuda*, which inhabits the Virgin River, a tributary of the Colorado River (USA). This hybrid species contacts but does not overlap its parent species, *Gila robusta* and *Gila elegans*, from the Colorado River. *Gila seminuda* is morphologically and genetically intermediate, and similar to synthetic hybrids. Intermediacy may allow it to out-compete both parents in the Virgin River[35]. A similar case, with well-documented genetic intermediacy, is an unnamed form of the butterfly genus *Lycaeides*. This homoploid hybrid species uses a different host plant and inhabits high-elevation alpine habitats unoccupied by either parent[36].

Rhagoletis fruitflies provide another historically documented example. In 1997 flies were first found on introduced honeysuckle (*Lonicera* spp.). Molecular markers in the fly are a blend from the two parents: the blueberry maggot *Rhagoletis mendax* and the snowberry maggot *Rhagoletis zephyria*[37]. No F_1 genotypes are detected between the hybrid form and its parents where they overlap; the new fly is reproductively isolated. *Rhagoletis* flies mate on host fruits, and host choice ensures mating specificity. A different route to hybrid speciation can be inferred from the Colombian butterfly *Heliconius heurippa*[38]. This form has a colour pattern like that of synthetic hybrids between local *Heliconius cydno* and *Heliconius melpomene*. Microsatellite alleles are shared across all three species, but *H. heurippa* forms an allele frequency cluster distinguishable from either parent. The hybrid wing coloration of *H. heurippa* is a cue in mating discrimination, and directly causes reproductive isolation from both parents. Similarly, in the fish *Xiphophorus clemenciae*, the 'swordtail', a hybridization-derived trait, is involved in sexual selection and mate choice[39] and may be related to its speciation.

Hybrid speciation in animals is supported so far only by low-resolution molecular data. Genomic mapping of ecological or speciation-related hybrid traits, which so strongly supports hybrid speciation in *Helianthus*, is not yet available for any animal case. Many homoploid hybrid species fail to overlap with at least one parent species, and reproductive isolation is weak, so species status could be questioned (the *Lonicera*-feeding *Rhagoletis* is an exception). Nonetheless, in the cases surveyed here, hybrid traits often contribute strongly to maintenance or ecological expansion of the new form.

Is hybrid speciation important in evolution?

There are now many examples of hybrid species. We know that polyploidy is common in plants, giving rise to ≥2–7% of vascular plant species, but rarer in animals. Furthermore, ancient polyploidy has been found at the root of many plant and animal groups. Genome duplications probably facilitated the evolution of complex organisms (although this is debated)[21], and we can infer that successful genome duplications were mostly allopolyploid, provided that limited plant community data are reliable[4,27]. Hybridization would then be a catalyst not only for speciation but also for major evolutionary innovations.

Polyploid speciation leaves a clear genomic signature, but we have little idea how common homoploid hybrid species are. They could be abundant: most speciation involves natural selection[6]; natural selection requires genetic variation; genetic variation is enhanced by hybridization[12]; and hybridization and introgression between species is a regular occurrence, especially in rapidly radiating groups[9,12,18]. Enough suspected animal homoploid hybrid species exist to indicate that it may be at least as common as in plants, in contrast to the situation for polyploidy, where a variety of traits prevent its occurrence (see above). It now seems intuitively unlikely that all biodiversity arose as a result of recombination of existing diversity[1], but homoploid hybrid species might still represent a large fraction. Nonetheless, there are few convincing cases, probably, in part, because of the difficulty of demonstrating that hybridization has led to speciation. We clearly need more genomic analyses. As for hybrid species as a whole, we have observed recent speciation in the laboratory or nature in seven genera discussed here (*Helianthus*, *Senecio*, *Primula*, *Spartina*, *Rhagoletis*, *Bombyx* and *Cottus*), and there are many other cases. It would be hard to find another mode of speciation so readily documented historically and so amenable to experimentation.

That hybrid species exist at all reveals something perhaps unexpected about adaptive landscapes. If hybrid 'hopeful monsters', with all their problems, are ever to survive in competition with their parents, they must be able to hit (and for polyploid species, hit almost exactly) new adaptive combinations of genes (Fig. 1). This implies both that many adaptive peaks are scattered about in the adaptive landscape, and also that many are unoccupied. Liberal adaptive landscapes are further supported by the successes of many introduced species, and by fossil evidence: for insects, angiosperms and many other groups, diversity seems to have been increasing more or less continuously over geological time[40].

The ability of hybrid species to invade hitherto unoccupied niches also means that hybridization can contribute to adaptive radiations such as African cichlid fish and Darwin's finches[7,9,12]. This principle is well demonstrated by the 'domestication niche'. Humans have unwittingly created many allopolyploid and other hybrid crops and domestic animals while selecting for transgressively high yields[4,7]. Even our own species may have a hybrid genomic ancestry[41,42], although this is contested[43]. Whichever way the debate about humans is resolved, it would be hardly surprising if hybridization was one trigger for the origin of *Homo sapiens*, the most invasive mammal on the planet[42].

1. Lotsy, J. P. *Evolution by Means of Hybridization* (Martinus Nijhoff, The Hague, 1916).
2. Mayr, E. *Animal Species and Evolution* (Harvard Univ. Press, Cambridge, Massachusetts, 1963).
3. Stebbins, G. L. *Processes of Organic Evolution* (Prentice-Hall, Englewood Cliffs, New Jersey, 1971).
4. Grant, V. *Plant Speciation* (Columbia Univ. Press, New York, 1981).
5. Barton, N. H. The role of hybridization in evolution. *Mol. Ecol.* **10**, 551–568 (2001). Medline
6. Coyne, J. A. & Orr, H. A. *Speciation* (Sinauer Associates, Sunderland, Massachusetts, 2004).
7. Anderson, E. & Stebbins, G. L. Hybridization as an evolutionary stimulus. *Evolution* **8**, 378–388 (1954).
8. Arnold, M. L. *Natural Hybridization and Evolution* (Oxford Univ. Press, Oxford, 1997).
9. Seehausen, O. Hybridization and adaptive radiation. *Trends Ecol. Evol.* **19**, 198–207 (2004).
10. Dowling, T. E. & Secor, C. L. The role of hybridization and introgression in the diversification of animals. *Annu. Rev. Ecol. Syst.* **28**, 593–620 (1997).
11. Bullini, L. Origin and evolution of animal hybrid species. *Trends Ecol. Evol.* **9**, 422–426 (1994).
12. Grant, P. R., Grant, B. R. & Petren, K. Hybridization in the recent past. *Am. Nat.* **166**, 56–57 (2005).
13. Mallet, J. A species definition for the modern synthesis. *Trends Ecol. Evol.* **10**, 294–299 (1995).
14. Butlin, R. Speciation by reinforcement. *Trends Ecol. Evol.* **2**, 8–12 (1987).
15. Ortíz-Barrientos, D., Counterman, B. A. & Noor, M. A. F. The genetics of speciation by reinforcement. *PLoS Biol.* **2**, e416 (2004).
16. Tunner, H. G. & Nopp, H. Heterosis in the common European water frog. *Naturwissenschaften* **66**, 268–269 (1979).
17. Fisher, R. A. *The Genetical Theory of Natural Selection* (Clarendon Press, Oxford, 1930).

18. Mallet, J. Hybridization as an invasion of the genome. *Trends Ecol. Evol.* **20**, 229–237 (2005).
19. Husband, B. C. Constraints on polyploid evolution: a test of the minority cytotype exclusion principle. *Proc. R. Soc. Lond. B* **267,** 217–223 (2000).
20. Wright, S. The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proc. XI Int. Congr. Genet. Hague* **1,** 356–366 (1932).
21. Otto, S. P. & Whitton, J. Polyploid incidence and evolution. *Annu. Rev. Genet.* 34, 401–437 (2000).
22. Ramsey, J. & Schemske, D. W. Neopolyploidy in flowering plants. *Annu. Rev. Ecol. Syst.* **33,** 589–639 (2002).
23. Astaurov, B. L. Experimental polyploidy in animals. *Annu. Rev. Genet.* **3,** 99–126 (1969).
24. Muller, H. J. Why polyploidy is rarer in animals than in plants. *Am. Nat.* **59,** 346–353 (1925).
25. Mable, B. K. 'Why polyploidy is rarer in animals than in plants': myths and mechanisms. *Biol. J. Linn. Soc.* **82,** 453–466 (2004).
26. Soltis, D. E., Soltis, P. S. & Tate, J. A. Advances in the study of polyploidy since plant speciation. *New Phytol.* **161,** 173–191 (2004).
27. Brochmann, C. *et al.* Polyploidy in arctic plants. *Biol. J. Linn. Soc.* **82,** 521–536 (2004).
28. Abbott, R. J. & Lowe, A. J. Origins, establishment and evolution of new polyploid species: *Senecio cambrensis* and *S. eboracensis* in the British Isles. *Biol. J. Linn. Soc.* **82,** 467–474 (2004).
29. Ainouche, M. L., Baumel, A. & Salmon, A. *Spartina anglica* C. E. Hubbard: a natural model system for analysing early evolutionary changes that affect allopolyploid genomes. *Biol. J. Linn. Soc.* **82,** 475–484 (2004).
30. Buerkle, C. A., Morris, R. J., Asmussen, M. A. & Rieseberg, L. H. The likelihood of homoploid hybrid speciation. *Heredity* **84**, 441–451 (2000).
31. Rieseberg, L. H. Hybrid origins of plant species. *Annu. Rev. Ecol. Syst.* **28,** 359–389 (1997).
32. Rieseberg, L. H., Raymond, O., Rosenthal, D. M., Lai, Z. & Livingstone, K. Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* **301**, 1211–1216 (2003).
33. Gross, B. L. & Rieseberg, L. H. The ecological genetics of homoploid hybrid speciation. *J. Hered.* **96**, 241–252 (2005).
34. Nolte, A. W., Freyhof, J., Stemshorn, K. C. & Tautz, D. An invasive lineage of sculpins, *Cottus* sp. (Pisces, Teleostei) in the Rhine with new habitat adaptations has originated from hybridization between old phylogeographic groups. *Proc. R. Soc. Lond. B* **272,** 2379–2387 (2005).
35. DeMarais, B. D., Dowling, T. E., Douglas, M. E., Minckley, W. L. & Marsh, P. C. Origin of *Gila seminuda* (Teleostei: Cyprinidae) through introgressive hybridization: implications for evolution and conservation. *Proc. Natl Acad. Sci. USA* **89**, 2747–2751 (1992).
36. Gompert, Z., Fordyce, J. A., Forister, M., Shapiro, A. M. & Nice, C. C. Homoploid hybrid speciation in an extreme habitat. *Science* **314**, 1923–1925 (2006).
37. Schwarz, D., Matta, B. M., Shakir-Botteri, N. L. & McPheron, B. A. Host shift to an invasive plant triggers rapid animal hybrid speciation. *Nature* **436**, 546–549 (2005).
38. Mavárez, J. *et al.* Speciation by hybridization in *Heliconius* butterflies. *Nature* **441**, 868–871 (2006).
39. Meyer, A., Salzburger, W. & Schartl, M. Hybrid origin of a swordtail species (Teleostei: *Xiphophorus clemenciae*) driven by sexual selection. *Mol. Ecol.* **15**, 721–730 (2006).
40. Labandeira, C. C. & Sepkoski, J. J. Insect diversity in the fossil record. *Science* **261**, 310–315 (1993).
41. Patterson, N., Richter, D. J., Gnerre, S., Lander, E. S. & Reich, D. Genetic evidence for complex speciation of humans and chimpanzees. *Nature* **441**, 1103–1108 (2006).
42. Evans, P. D., Mekel-Bobrov, N., Vallender, E. J., Hudson, R. R. & Lahn, B. T. Evidence that the adaptive allele of the brain size gene microcephalin introgressed into *Homo sapiens* from an archaic *Homo* lineage. *Proc. Natl Acad. Sci. USA* **103**, 18178–18183 (2006).
43. Barton, N. H. How did the human species form? *Curr. Biol.* **16**, R647–R650 (2006).
44. Seehausen, O., van Alphen, J. J. M. & Witte, F. Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science* **277,** 1808–1811 (1997).
45. Grant, B. R. & Grant, P. R. High survival of Darwin's finch hybrids — effects of beak morphology and diets. *Ecology* **77,** 500–509 (1996).

Acknowledgements I thank C. Brochmann, K. Dasmahapatra, B. Husband, S. Knapp, M. Linares, J. Mavárez, A. Meyer, P. Nosil, S. Otto, C. Salazar and S. Turelli for discussions and comments. The work was supported, in part, by grants from NERC and the DEFRA Darwin Initiative programme.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The author declares no competing financial interests. Correspondence should be addressed to J.M. (j.mallet@ucl.ac.uk).

PROGRESS

The molecular choreography of a store-operated calcium channel

Richard S. Lewis[1]

Store-operated calcium channels (SOCs) serve essential functions from secretion and motility to gene expression and cell growth. A fundamental mystery is how the depletion of Ca^{2+} from the endoplasmic reticulum (ER) activates Ca^{2+} entry through SOCs in the plasma membrane. Recent studies using genetic approaches have identified genes encoding the ER Ca^{2+} sensor and a prototypic SOC, the Ca^{2+}-release-activated Ca^{2+} (CRAC) channel. New findings reveal a unique mechanism for channel activation, in which the CRAC channel and its sensor migrate independently to closely apposed sites of interaction in the ER and the plasma membrane.

Cell-surface receptors that activate phospholipase C (PLC) create cytosolic Ca^{2+} signals from intra- and extra-cellular sources. An initial transient rise of $[Ca^{2+}]_i$ (intracellular calcium concentration) results from the release of Ca^{2+} from the ER, which has been understood since the early 1980s to be triggered by the PLC product, inositol-1,4,5-trisphosphate ($InsP_3$), opening $InsP_3$ receptors in the ER[1]. However, a subsequent phase of sustained Ca^{2+} entry across the plasma membrane has long eluded a mechanistic description. On the basis of relationships between store emptying, Ca^{2+} entry and store refilling in parotid gland acinar cells, Putney proposed in 1986 a 'capacitative Ca^{2+} entry' hypothesis in which the emptying of Ca^{2+} stores itself activates Ca^{2+} channels in the plasma membrane to help refill the stores[2]. Over the years, this hypothesis (later renamed store-operated Ca^{2+} entry, or SOCE) has received widespread support from Ca^{2+}-imaging experiments in many types of excitable and non-excitable cells, in which store depletion by thapsigargin and other inhibitors of ER Ca^{2+} uptake were shown to activate Ca^{2+} entry independently of surface receptors or $InsP_3$ (ref. 3). Importantly, these studies revealed that SOCE does more than simply provide Ca^{2+} for refilling stores, but can itself generate sustained Ca^{2+} signals that control such essential functions as gene expression, cell metabolism and exocytosis[4].

But characterizing SOCE pathways using Ca^{2+}-imaging techniques alone is problematic; cytosolic Ca^{2+} signals reflect the combined activities of channels, pumps, transporters, intracellular buffers, membrane potential and organelles, all of which can interact through changes in $[Ca^{2+}]_i$. Thus, a crucial step in establishing a specific molecular basis for SOCE was the use of the patch-clamp to identify and characterize a store-operated Ca^{2+} current, named the Ca^{2+}-release-activated Ca^{2+} (CRAC) current, or I_{CRAC}, in mast cells[5] and Jurkat leukaemic T cells[6,7]. Extensive patch-clamp studies established the CRAC entry mechanism as an ion channel with distinctive biophysical characteristics, including an extremely high selectivity for Ca^{2+} paired with an exceedingly low conductance[4,8]. Most importantly, the CRAC channel was shown to fulfil the rigorous criteria for being store-operated: activation solely by the reduction of Ca^{2+} in the ER rather than by cytosolic Ca^{2+} or other messengers generated by PLC[8] (Box 1). Clinical studies demonstrate that the CRAC channel is absolutely required for the activation of genes underlying the T cell response to antigen[9,10]. Although a variety of store-operated channels (SOCs) may exist[4] (see below), the CRAC channel's distinctive biophysical fingerprint, quantifiable store-dependence and essential function in T cells has made it the prototypic SOC and the most sought-after for cloning by molecular cell biologists[8].

Years of frustration marked the quest for a molecular basis of SOCE. Biochemical approaches were stymied by the lack of high-affinity ligands or an abundant source of SOCs, and molecular cloning was thwarted by the presence of endogenous SOCE in most expression systems and the lack of significant functional homology of CRAC to the known Ca^{2+} channels. Candidate gene approaches, mostly focusing on the TRP channel family, were unproductive, in large part because of a lack of consistent evidence showing that they are store-operated (Box 1), and because their pore properties differ significantly from those of CRAC[4,8]. Against this backdrop, a number of models for SOCE were proposed, but could not be definitively tested in the absence of identified molecular components[4,8].

This era of uncertainty ended recently with the successful application of forward and reverse genetic approaches. In 2005, STIM1 (stromal interaction molecule 1) was identified as the mammalian ER Ca^{2+} sensor[11,12], closely followed in 2006 by the identification of Orai1/CRACM1 as a component of the mammalian CRAC channel[13–15]. These breakthrough studies and the flurry of publications that followed have stoked rapid progress in illuminating key steps in the molecular mechanism of SOCE.

STIM1 is the ER Ca^{2+} sensor for store-operated Ca^{2+} entry

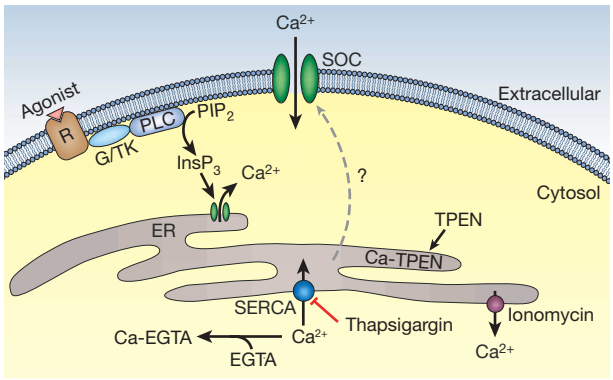
Two groups identified STIM1 in RNA-mediated interference (RNAi) screens for suppressors of SOCE in *Drosophila* S2 cells[12] and HeLa cells[11]. STIM1 is a 77 kDa type I membrane protein with multiple predicted protein interaction or signalling domains (Fig. 1), and is located predominantly in the ER[11,16,17] but also to a limited extent in the plasma membrane[17–19]. Knockdown of STIM1 in mammalian cells (or the *Drosophila* homologue in S2 cells) significantly reduces SOCE and I_{CRAC}[11,12,20–22]. Its widespread tissue distribution and conservation from *Drosophila* to human is consistent with the prevalence of SOCE among cells and organisms.

Total internal reflection fluorescence (TIRF) and confocal microscopy reveal that STIM1 is distributed throughout the ER when Ca^{2+} stores are full, but redistributes into discrete puncta near the plasma membrane on store depletion[11,16,17,19,21,23–25]. The exact location of these puncta has been a matter of some debate. One study suggested that store depletion causes the insertion of STIM1 into the plasma membrane

[1]Department of Molecular and Cellular Physiology, Stanford University School of Medicine, Stanford, California 94305, USA.

Box 1 | The defining characteristics of store-operated channels

Experimentally, SOCs are strictly defined by their response to agents that directly deplete Ca^{2+} stores, such as SERCA pump inhibitors (for example, thapsigargin), Ca^{2+} ionophores (ionomycin), cytosolic Ca^{2+} buffers (EGTA), and ER Ca^{2+} buffers (TPEN, *N,N,N′,N′*-tetrakis (2-pyridylmethyl)ethylene diamine). To measure directly SOC activity without contamination from channels sensitive to cytosolic Ca^{2+}, these tests should be conducted by patch-clamp recording in the whole-cell or perforated-patch configuration with a high level of intracellular buffering to keep $[Ca^{2+}]_i$ constant. The CRAC channel has been consistently shown to be activated under all of these conditions, and store depletion is its only known activator[8]. In contrast, although TRPC1 is often referred to as a SOC, results of patch-clamp studies are divided on the answer to whether it is store-operated[4,40], and it can also be activated by diacylglycerol[41] and membrane stretch[42]. Rigorous assays outlined above will be needed to assess its status as a true SOC.



Box 1 Figure 1 | Physiological pathway of SOC activation. Agonist binding to receptors (R) activates phospholipase C (PLC) through a G protein or tyrosine kinase. PLC hydrolyses PIP_2 in the plasma membrane to generate $InsP_3$, which releases ER Ca^{2+} through $InsP_3$ receptors. The resulting reduction of ER $[Ca^{2+}]$ activates SOCs (dashed line).

where it may control SOCE[17]; this conclusion was based primarily on an increase in the amount of STIM1 pulled down with biotinylated plasma membrane proteins after depletion (but see ref. 19). Biotinylation of STIM1 itself was not confirmed, however, allowing the possibility that ER-localized STIM1 may have been isolated by association with a protein in the plasma membrane (see below). On the other hand, an intracellular location for puncta is suggested by studies of cells expressing STIM1 tagged with peptides or fluorescent proteins, all of which failed to detect extracellular exposure of the tag, which would be expected from STIM1 insertion in the plasma membrane[11,16,21,24,25]. The most precise localization has been achieved by electron microscopy of cells expressing a STIM1–horseradish-peroxidase (HRP) fusion protein, showing that the puncta represent STIM1 accumulation within junctional ER structures located 10–25 nm from the plasma membrane[16]. Although protein tags might perturb STIM1 trafficking and prevent its insertion into the plasma membrane, the tagged proteins do seem to be functional in activating SOCE[11,16, 21,25]; thus, the significance of endogenous STIM1 insertion, if it does occur, is as yet unclear.

Although the redistribution of STIM1 into junctional ER regions is slow[11,17], it does precede the opening of CRAC channels by several seconds[16] and is therefore rapid enough to be an essential step in the activation of CRAC channels. The critical evidence for STIM1 as the Ca^{2+} sensor for SOCE is that mutation of predicted Ca^{2+}-binding residues of the EF hand structural motif, expected to reduce its affinity for Ca^{2+} and hence mimic the store-depleted state, causes STIM1 to redistribute spontaneously into puncta and trigger constitutive Ca^{2+} influx through SOCs even when stores are full[11,17,21,22]. The affinity of the isolated EF hand *in vitro* is 200–600 μM[26], consistent with a Ca^{2+}-sensing function given the range of $[Ca^{2+}]$ reported for the ER lumen (250–600 μM)[27].

Cloning the elusive CRAC channel

Remarkably, after more than a decade of attempts to identify the CRAC channel gene, three groups converged on it within a few months of each other. The first approach began over 6 yr ago with studies of T cells from human patients with a severe combined immunodeficiency (SCID) syndrome, in which T cell receptor engagement or store depletion failed to activate Ca^{2+} entry[10,28]. Linkage analysis and positional cloning led to a region of chromosome 12 containing ~74 genes. Meanwhile, a genome-wide RNAi screen in S2 cells for blockers of the nuclear translocation of NFAT, a Ca^{2+}-dependent transcription factor, yielded one hit having a human homologue within the targeted region of chromosome 12, a previously uncharacterized gene that was subsequently named *Orai1* (ref. 13). Two other screens of the same RNAi library for suppression of SOCE identified the same gene in S2 cells[14,15], and the human homologue was given a second name, *CRACM1*[14]. Orai1/CRACM1 is a widely expressed, 33 kDa plasma membrane protein with 4 transmembrane domains and a lack of significant sequence homology to other ion channels (Fig. 1). The human SCID defect was traced to a single point mutation in *Orai1* (Fig. 1); importantly, retroviral transduction with the wild-type gene rescued the Ca^{2+}-signalling defect, showing unequivocally that Orai1 is essential for CRAC channel activity[13]. Functions of the two other mammalian Orai homologues, Orai2 and Orai3, are not as clear; although they can exhibit SOC activity when overexpressed with STIM1 in HEK cells[21], moderate retroviral expression of Orai2 and Orai3 fails to rescue Ca^{2+} influx in

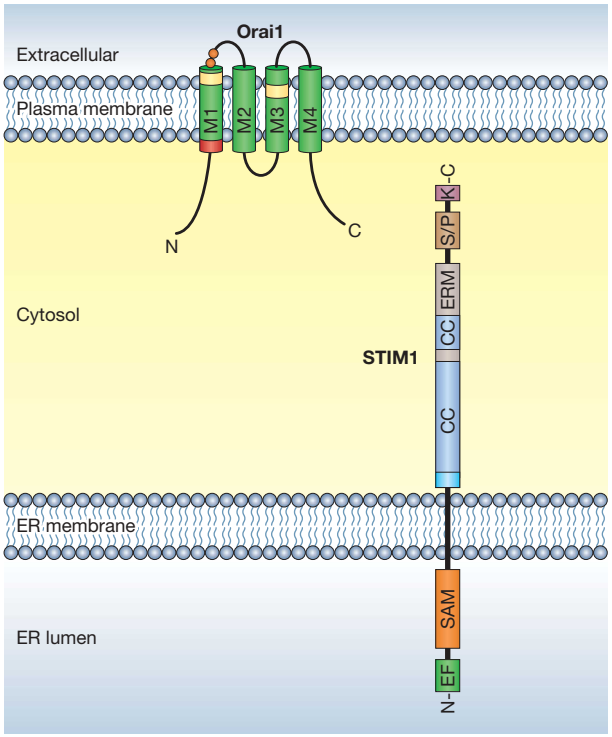


Figure 1 | STIM1 senses ER Ca^{2+} to activate Orai1, a pore-forming subunit of the CRAC channel. STIM1 is localized primarily in the ER membrane, with a small fraction in the plasma membrane (not pictured). The organization of the major predicted domains is shown, including an unpaired EF hand and sterile-α motif (SAM) domains on the luminal side, and on the cytosolic side, overlapping coiled-coil (CC) and ezrin-radixin-moesin (ERM) domains, and serine-proline-rich (S/P) and lysine-rich (K) domains. Orai1 is a plasma membrane protein with four membrane-spanning regions and intracellular N and C termini. Several critical residues are highlighted. A spontaneous mutation (R91W; red) causes a loss of CRAC channel function and immunodeficiency in human patients[13]. Glutamates at positions 106 and 190 (yellow) and aspartates at positions 110 and 112 (orange) affect ion selectivity, permeation and block[32–34].

human SCID T cells expressing the inactive mutant Orai1 (S. Feske, S. Srikanth and A. Rao, personal communication).

But is Orai1 part of the CRAC channel itself? Although overexpression of Orai1 together with STIM1 (or the same with the *Drosophila* homologues) was found to produce enormous CRAC currents[15,21,29,30], this result did not exclude the possibility that *Orai1* encodes a necessary non-channel component or regulator that is normally present in limiting amounts. Thus, three groups carried out a more definitive test, asking whether mutations can affect an intrinsic pore property such as ion selectivity. Earlier patch-clamp studies[4] suggested that CRAC channels attain high Ca^{2+} selectivity by a mechanism originally described for voltage-gated Ca^{2+} (Ca_V) channels, in which Ca^{2+} binds to conserved acidic pore residues to block permeation of monovalent cations[31]. *Orai1* mutagenesis revealed two glutamates to be particularly critical (Fig. 1). Substitution of aspartate at Glu 106 weakens the ability of Ca^{2+} to block Na^+ flux and reduces the channel's Ca^{2+} selectivity, whereas it increases permeability to Ba^{2+} and Cs^+ (refs 32–34). A glutamine substitution at Glu 190 also reduces Ca^{2+} selectivity and increases Cs^+ permeability[33]. Finally, neutralizing the charge on a pair of aspartates in the I–II loop (Asp 110 and Asp 112) reduces block by Gd^{3+} (ref. 32) and block of outward current by extracellular Ca^{2+} (ref. 33), indicating that these negatively charged sites may promote accumulation of polyvalent cations near the mouth of the pore. Besides providing the first glimpse into how CRAC channels attain their extreme selectivity for Ca^{2+}, these studies offer definitive evidence that Orai1 contributes to the CRAC channel pore. Given that currents through overexpressed Orai1 closely resemble I_{CRAC}, and that Orai1 can form multimers[32–34], it seems likely that the native CRAC channel is either a multimer of Orai1 alone or in combination with the closely related subunits Orai2 and/or Orai3. Thus, it seems that after 20 yr, the Ca^{2+} sensor and its target, the CRAC channel, have finally been found.

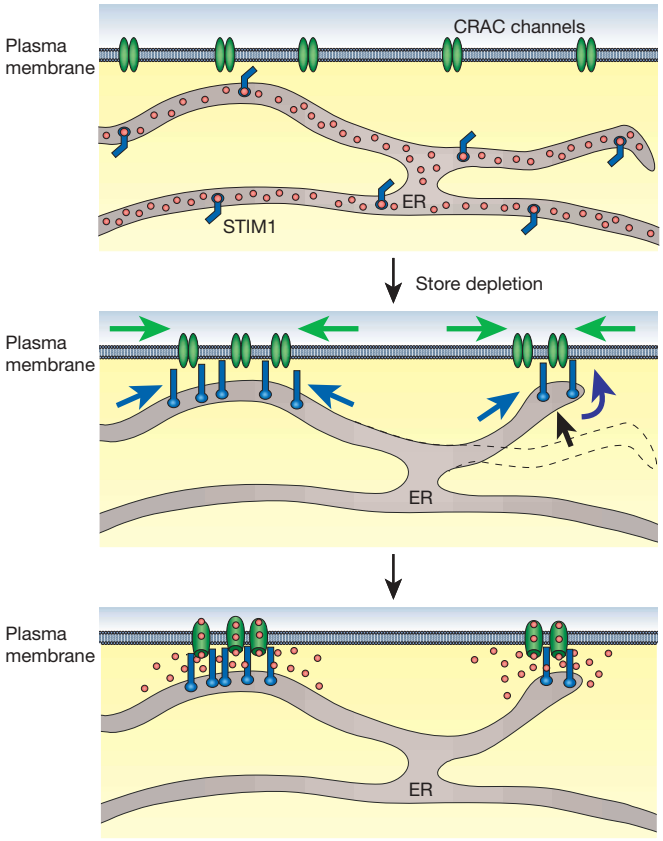


Figure 2 | The functional units of store-operated Ca^{2+} entry assemble in response to store depletion. In resting cells with replete Ca^{2+} stores (top panel), STIM1 and Orai1 are dispersed throughout the ER and plasma membrane, respectively. Store depletion (middle panel) causes STIM1 to accumulate at locations where the ER is juxtaposed to the plasma membrane ('junctional ER'), and to a limited extent increases the number of these close contacts (black arrow)[16]. At the same time, Orai1 accumulates in regions of the plasma membrane directly opposite the STIM1 clusters. The co-localization of STIM1 and Orai1 restricts channel activation and Ca^{2+} entry to these sites (bottom panel)[23]. Adapted with permission, from refs 23 and 39.

The molecular choreography of store-operated Ca^{2+} entry

STIM1 and Orai1 provide the first molecular tools for addressing the fundamental question underlying SOCE: how is the loss of Ca^{2+} from the ER lumen coupled to the opening of CRAC channels in the plasma membrane? As described above, depletion is sensed by STIM1, causing it to accumulate in junctional ER adjacent to the plasma membrane. In a TIRF-based Ca^{2+}-imaging study to map the locations of open CRAC channels, $[Ca^{2+}]_i$ elevations were seen to co-localize with STIM1 puncta, showing directly that CRAC channels open only in extreme proximity to these sites[23]. How is such local communication between STIM1 and the CRAC channel achieved? Remarkably, in cells co-expressing labelled STIM1 and Orai1, store depletion causes Orai1 itself to move from a dispersed distribution to accumulate in the plasma membrane directly opposite STIM1, presumably enabling STIM1 to activate the channel[23,24] (Fig. 2).

These studies describe for the first time the elementary unit of SOCE: apposed clusters of STIM1 in the ER and Orai1 in the plasma membrane, separated by a narrow gap of cytosol[16,23,24]. These structures are reminiscent of the functional units of Ca^{2+} release in striated muscle, where dihydropyridine receptors in the plasma membrane trigger Ca^{2+} release from ryanodine receptors in the closely apposed sarcoplasmic reticulum[35]. However, one important difference is that functional couplings in muscle are 'hard-wired' with channels in place before the stimulus arrives (ensuring rapid and reliable excitation–contraction coupling), whereas SOCE couplings are assembled on demand. This kind of choreographic activation mechanism, in which a channel and its sensor migrate within distinct membranes to reach a common interaction site, is unprecedented. By introducing delays between changes in store content and channel activity, it may promote fluctuations in Ca^{2+} entry that have been linked to Ca^{2+} oscillations in T cells[6,36]. Also, the confinement of SOCE to junctional sites might enhance the specificity of Ca^{2+} signalling by creating local signalling domains, perhaps explaining how signalling proteins such as adenylate cyclase, nitric oxide synthase, Cl^- channels and Ca^{2+}-ATPases (see ref. 4 for references) can be preferentially activated through SOCE.

A key question is how does STIM1 activate Orai1 at the ER–plasma-membrane junction? The junctional gap (10–25 nm) may be small enough to permit protein–protein interactions. In fact, overexpressed STIM1 and Orai1 can be co-immunoprecipitated[32,33], and the amount may increase following store depletion[32], suggesting that STIM1 and Orai1 interact either directly or as members of a multiprotein complex. Consistent with this idea, the expression of the cytosolic portion of STIM1 by itself was sufficient to activate CRAC channels in one study[20], and the effects of deleting the ERM/coiled-coil and other C-terminal domains suggest roles in STIM1 clustering and SOC activation[20,25]. On the luminal side of STIM1, the isolated EF-SAM region forms dimers and higher-order multimers on removal of Ca^{2+} *in vitro*, indicating that STIM1 oligomerization may be an early step in SOC activation[26].

Future directions

Even as a molecular picture of CRAC channel activation and SOCE is beginning to emerge, a multitude of new questions arise. How is the movement of STIM1 and Orai1 controlled (active or passive), and what molecular cues cause them to accumulate at junctional sites? After their arrival, how does STIM1 trigger CRAC channel opening? Although a physical interaction seems most likely (see above), the local generation of a diffusible activator[37] has not been ruled out. Biochemical and structural analysis of STIM1 and Orai1

will be key to understanding how the Ca^{2+} sensor and the channel communicate.

The function of STIM1 in the plasma membrane is enigmatic. Antibodies to the extracellular domain of STIM1 inhibit I_{CRAC}[22], suggesting a requisite role, but labelled STIM1 proteins confined to the ER can restore SOCE after endogenous STIM1 has been knocked down or out[11,21,25]. The function of STIM2 is also unclear: knockdown has been reported to have no effect[12] or an inhibitory effect on SOCE[11], whereas STIM2 overexpression is reported to inhibit SOCE or constitutively activate overexpressed Orai1 (ref. 30). It remains to be shown whether these discrepancies reflect differences in cell types, expression levels, or other factors.

The identification of Orai1 finally opens the door to studying the structural basis of CRAC channel behaviour. The unusual characteristics of this channel have long intrigued ion-channel biophysicists; it selects for Ca^{2+} just as well as Ca_V channels but conducts Ca^{2+} >100 times more slowly, is inactivated by intracellular Ca^{2+} on timescales separated by three orders of magnitude, and requires extracellular Ca^{2+} to be fully active[4,8]. Determining the structure of Orai1 will be essential for guiding mutagenesis-based studies of these unique channel properties.

Finally, one must ask, do the *Orai* genes account for all SOCs, or are there more SOC genes waiting to be discovered? The T cells and S2 cells used to isolate the *Orai* genes seem to express CRAC as their only SOC [9,28,38], so it is perhaps no surprise that other SOC genes did not emerge. There is evidence that SOCs distinct from CRAC exist (for example, in salivary gland, epidermal, endothelial, smooth muscle and pancreatic acinar cells[4,8]). Interestingly, although the R91W mutation in Orai1 completely blocks SOCE in human T cells, it only partially suppresses influx in EBV-transformed B cells, and a patient carrying this mutation did not show widespread deficits in extra-immune organs that are known to express SOCs[13]. These observations could be explained by tissue-specific expression of other SOC subtypes; alternatively, in some tissues store-independent Ca^{2+}-signalling pathways may be functionally redundant with SOCE, or SOCE itself may serve a contributory rather than an essential role (to date, the T cell is the only example where SOCE has been shown to be absolutely required for a critical cell behaviour). Future studies of SOCE in *Orai*-knockout animals will help to clarify the relation of Orai to SOCE in different tissues, and ultimately provide a more global view of the molecular workings of store-operated channels and their physiological roles.

1. Streb, H., Irvine, R. F., Berridge, M. J. & Schulz, I. Release of Ca^{2+} from a nonmitochondrial intracellular store in pancreatic acinar cells by inositol-1,4,5-trisphosphate. *Nature* **306,** 67–69 (1983).
2. Putney, J. W. Jr. A model for receptor-regulated calcium entry. *Cell Calcium* **7,** 1–12 (1986).
3. Putney, J. W. Jr & Bird, G. S. The inositol phosphate–calcium signaling system in nonexcitable cells. *Endocr. Rev.* **14,** 610–631 (1993).
4. Parekh, A. B. & Putney, J. W. Jr. Store-operated calcium channels. *Physiol. Rev.* **85,** 757–810 (2005).
5. Hoth, M. & Penner, R. Depletion of intracellular calcium stores activates a calcium current in mast cells. *Nature* **355,** 353–356 (1992).
6. Lewis, R. S. & Cahalan, M. D. Mitogen-induced oscillations of cytosolic Ca^{2+} and transmembrane Ca^{2+} current in human leukemic T cells. *Cell Regul.* **1,** 99–112 (1989).
7. Zweifach, A. & Lewis, R. S. Mitogen-regulated Ca^{2+} current of T lymphocytes is activated by depletion of intracellular Ca^{2+} stores. *Proc. Natl Acad. Sci. USA* **90,** 6295–6299 (1993).
8. Prakriya, M. & Lewis, R. S. Store-operated calcium channels: properties, functions and the search for a molecular mechanism. In *Molecular and Cellular Insights into Ion Channel Biology* (ed. Robert Maue) 121–140 (Elsevier Science, Amsterdam, 2004).
9. Partiseti, M. *et al.* The calcium current activated by T cell receptor and store depletion in human lymphocytes is absent in a primary immunodeficiency. *J. Biol. Chem.* **269,** 32327–32335 (1994).
10. Feske, S., Giltnane, J., Dolmetsch, R., Staudt, L. M. & Rao, A. Gene regulation mediated by calcium signals in T lymphocytes. *Nature Immunol.* **2,** 316–324 (2001).
11. Liou, J. *et al.* STIM is a Ca^{2+} sensor essential for Ca^{2+}-store-depletion-triggered Ca^{2+} influx. *Curr. Biol.* **15,** 1235–1241 (2005).
12. Roos, J. *et al.* STIM1, an essential and conserved component of store-operated Ca^{2+} channel function. *J. Cell Biol.* **169,** 435–445 (2005).
13. Feske, S. *et al.* A mutation in Orai1 causes immune deficiency by abrogating CRAC channel function. *Nature* **441,** 179–185 (2006).
14. Vig, M. *et al.* CRACM1 is a plasma membrane protein essential for store-operated Ca^{2+} entry. *Science* **312,** 1220–1223 (2006).
15. Zhang, S. L. *et al.* Genome-wide RNAi screen of Ca^{2+} influx identifies genes that regulate Ca^{2+} release-activated Ca^{2+} channel activity. *Proc. Natl Acad. Sci. USA* **103,** 9357–9362 (2006).
16. Wu, M. M., Buchanan, J., Luik, R. M. & Lewis, R. S. Ca^{2+} store depletion causes STIM1 to accumulate in ER regions closely associated with the plasma membrane. *J. Cell Biol.* **174,** 803–813 (2006).
17. Zhang, S. L. *et al.* STIM1 is a Ca^{2+} sensor that activates CRAC channels and migrates from the Ca^{2+} store to the plasma membrane. *Nature* **437,** 902–905 (2005).
18. Manji, S. S. *et al.* STIM1: a novel phosphoprotein located at the cell surface. *Biochim. Biophys. Acta* **1481,** 147–155 (2000).
19. Soboloff, J. *et al.* STIM2 is an inhibitor of STIM1-mediated store-operated Ca^{2+} entry. *Curr. Biol.* **16,** 1465–1470 (2006).
20. Huang, G. N. *et al.* STIM1 carboxyl-terminus activates native SOC, I_{crac} and TRPC1 channels. *Nature Cell Biol.* **8,** 1003–1010 (2006).
21. Mercer, J. C. *et al.* Large store-operated calcium selective currents due to co-expression of Orai1 or Orai2 with the intracellular calcium sensor, Stim1. *J. Biol. Chem.* **281,** 24979–24990 (2006).
22. Spassova, M. A. *et al.* STIM1 has a plasma membrane role in the activation of store-operated Ca^{2+} channels. *Proc. Natl Acad. Sci. USA* **103,** 4040–4045 (2006).
23. Luik, R. M., Wu, M. M., Buchanan, J. & Lewis, R. S. The elementary unit of store-operated Ca^{2+} entry: local activation of CRAC channels by STIM1 at ER–plasma membrane junctions. *J. Cell Biol.* **174,** 815–825 (2006).
24. Xu, P. *et al.* Aggregation of STIM1 underneath the plasma membrane induces clustering of Orai1. *Biochem. Biophys. Res. Commun.* **350,** 969–976 (2006).
25. Baba, Y. *et al.* Coupling of STIM1 to store-operated Ca^{2+} entry through its constitutive and inducible movement in the endoplasmic reticulum. *Proc. Natl Acad. Sci. USA* **103,** 16704–16709 (2006).
26. Stathopulos, P. B., Li, G. Y., Plevin, M. J., Ames, J. B. & Ikura, M. Stored Ca^{2+} depletion-induced oligomerization of STIM1 via the EF-SAM region: An initiation mechanism for capacitive Ca^{2+} entry. *J. Biol. Chem.* **281,** 35855–35862 (2006).
27. Demaurex, N. & Frieden, M. Measurements of the free luminal ER Ca^{2+} concentration with targeted "cameleon" fluorescent proteins. *Cell Calcium* **34,** 109–119 (2003).
28. Feske, S., Prakriya, M., Rao, A. & Lewis, R. S. A severe defect in CRAC Ca^{2+} channel activation and altered K^+ channel gating in T cells from immunodeficient patients. *J. Exp. Med.* **202,** 651–662 (2005).
29. Peinelt, C. *et al.* Amplification of CRAC current by STIM1 and CRACM1 (Orai1). *Nature Cell Biol.* **8,** 771–773 (2006).
30. Soboloff, J. *et al.* Orai1 and STIM reconstitute store-operated calcium channel function. *J. Biol. Chem.* **281,** 20661–20665 (2006).
31. Sather, W. A. & McCleskey, E. W. Permeation and selectivity in calcium channels. *Annu. Rev. Physiol.* **65,** 133–159 (2003).
32. Yeromin, A. V. *et al.* Molecular identification of the CRAC channel by altered ion selectivity in a mutant of Orai. *Nature* **443,** 226–229 (2006).
33. Vig, M. *et al.* CRACM1 multimers form the ion-selective pore of the CRAC channel. *Curr. Biol.* **16,** 2073–2079 (2006).
34. Prakriya, M. *et al.* Orai1 is an essential pore subunit of the CRAC channel. *Nature* **443,** 230–233 (2006).
35. Franzini-Armstrong, C., Protasi, F. & Tijskens, P. The assembly of calcium release units in cardiac muscle. *Ann. NY Acad. Sci.* **1047,** 76–85 (2005).
36. Dolmetsch, R. E. & Lewis, R. S. Signaling between intracellular Ca^{2+} stores and depletion-activated Ca^{2+} channels generates $[Ca^{2+}]_i$ oscillations in T lymphocytes. *J. Gen. Physiol.* **103,** 365–388 (1994).
37. Bolotina, V. M. & Csutora, P. CIF and other mysteries of the store-operated Ca^{2+}-entry pathway. *Trends Biochem. Sci.* **30,** 378–387 (2005).
38. Yeromin, A. V., Roos, J., Stauderman, K. A. & Cahalan, M. D. A store-operated calcium channel in *Drosophila* S2 cells. *J. Gen. Physiol.* **123,** 167–182 (2004).
39. Luik, R. M. & Lewis, R. S. New insights into the molecular mechanisms of store-operated Ca^{2+} signaling in T cells. *Trends Mol. Med.*, doi:10.1016/j.molmed.2007.01.004 (30 January 2007).
40. Strübing, C., Krapivinsky, G., Krapivinsky, L. & Clapham, D. E. TRPC1 and TRPC5 form a novel cation channel in mammalian brain. *Neuron* **29,** 645–655 (2001).
41. Lintschinger, B. *et al.* Coassembly of Trp1 and Trp3 proteins generates diacylglycerol- and Ca^{2+}-sensitive cation channels. *J. Biol. Chem.* **275,** 27799–27805 (2000).
42. Maroto, R. *et al.* TRPC1 forms the stretch-activated cation channel in vertebrate cells. *Nature Cell Biol.* **7,** 179–185 (2005).

Acknowledgements The author thanks M. Wu and R. Luik for stimulating discussions throughout the course of this work. This work was supported by a grant from the NIGMS and the Mathers Charitable Foundation.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The author declares no competing financial interests. Correspondence should be addressed to R.S.L. (rslewis@stanford.edu).

SUPERCONDUCTORS

Unusual oxygen isotope effects in cuprates?

Arising from: G.-H. Gweon *et al. Nature* 430, 187–190 (2004)

The possibility that a pairing boson might act as the 'glue' to bind electrons into a Cooper pair in superconductors with a high critical temperature (T_c) is being actively pursued in condensed-matter physics. Gweon *et al.*[1] claim that there is a large and unusual oxygen-isotope effect on the electronic structure, indicating that phonons have a special importance in high-temperature superconductors. However, we are unable to detect this unusual oxygen-isotope effect in new data collected under almost identical material and experimental conditions. Our findings point towards a more conventional influence of phonons in these materials.

In the high-temperature superconductors, oxygen-isotope exchange has been found to have a very small effect on the T_c, particularly for optimally doped (highest-T_c) samples[2]. It was therefore unexpected when Gweon *et al.* reported a large isotope effect on the electronic structure in optimally doped $Bi_2Sr_2CaCu_2O_{8+\delta}$ using angle-resolved photoemission spectroscopy (ARPES)[1]. This isotope effect was seen mainly at a broad, high-energy hump feature and was unusually large (10–40 meV shift), by contrast with the expected isotopic phonon shifts of the order of 3 meV. It was argued by many that this is experimental evidence that the coupling between the lattice and electrons plays a particularly significant role.

Figure 1a shows our energy–momentum (E–k) dispersion relations for both ^{16}O and ^{18}O taken along the 'nodal' direction of the Brillouin zone. The results show an insignificant shift for the deeper states of -0.9 ± 0.4 meV, by contrast with Fig. 1b of Gweon *et al.*[1], which shows a 15-meV shift of more deeply lying states. We found that precise rotational alignment of the samples was crucial to obtain repeatable energy positions, especially in the high-energy regime discussed here. The negligible energy shift was confirmed by using three different facilities, samples with multiple doping levels, including the optimal doping described in ref. 1, and with various photon energies, including the same energy used by Gweon *et al.*[1] (33 eV), and a low photon energy (7.0 eV) that gives extreme resolution, as well as increased bulk sensitivity.

Gweon *et al.*[1] show an even larger effect (up to 40 meV) at the 'antinodal' portions of the Brillouin zone (towards the $(0,\pi)$ and $(0,-\pi)$ points) — comparable data are shown here in Fig. 1b. We studied ten cuts that had a similar geometry to the key cut 7 of Gweon *et al.*[1], with half of the cuts at positive k_y (momentum along the y direction) values and the others at negative k_y. Having data on both sides of the (0,0) point allowed the correction of any very slight (0.1 degrees) sample misalignments, which is crucial for removing systematic alignment errors that could cause apparent energy shifts. Figure 1b (right) shows the dispersion of the bottom of the band (determined in each cut as in Fig. 2d of Gweon *et al.*[1]) along the $(0,-\pi)\rightarrow(0,\pi)$ symmetry line. The ten sets in Fig. 1b have an average isotope shift of the band bottom of 2 ± 3 meV, which is inconsistent with the 10–40 meV level found by Gweon *et al.*[1].

Our result is not inconsistent with the tunnelling result of a conventional phononic isotope shift on the scale of about 3 meV (refs 3,4) that was performed on the same samples and which conceivably could be seen near the kink energy in future ultra-resolution ARPES.

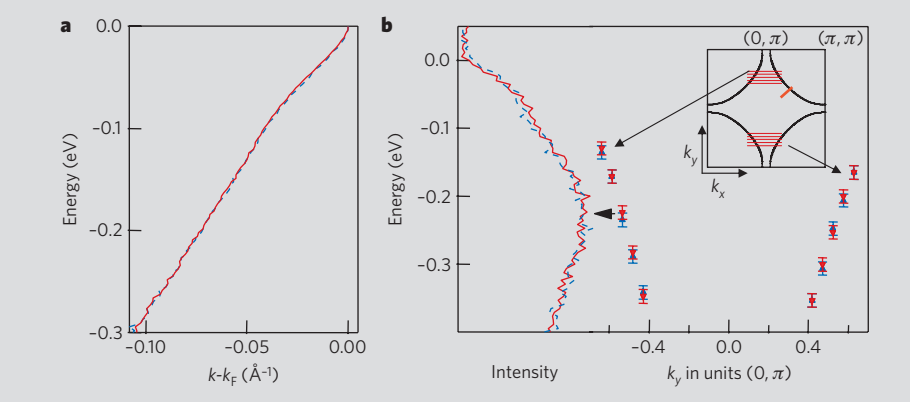


Figure 1 | Comparison of nodal and antinodal energy (E) and momentum (k) dispersions as a function of oxygen-isotope exchange. All data are from optimally doped (T_c = 92 K, unexchanged) $Bi_2Sr_2CaCu_2O_{8+\delta}$ crystals taken with 33 eV photons at 15 K. Blue line and symbols, ^{16}O data; red line and symbols, ^{18}O data. No large-scale shift of the spectra with isotope exchange is evident. **a**, Nodal dispersion (red diagonal line in **b**, inset) obtained by fitting momentum-distribution curves, in which k_F is the Fermi momentum. **b** (left), Energy-distribution curve from each isotope at the k-point indicated (arrowhead); right, dispersion of the band bottom from energy-distribution peaks along the $(0, -\pi)$ –$(0, \pi)$ symmetry line (horizontal lines in the Fermi surface inset).

Although we do not observe the unusual large-scale isotope effect reported by Gweon *et al.*[1], our result does not invalidate electron–phonon coupling as a potential pairing mechanism for high-temperature superconductivity. Rather, it implies that if electron–phonon coupling is responsible, it must do it in a more subtle, roundabout way.

Methods. Our data were measured with 33-eV photons at BL12 of the Advanced Light Source, Berkeley, USA, by use of a high-precision 6-axis sample manipulator. Single crystals of $Bi_{1.9}Sr_{2.1}CaCu_2O_{8+\delta}$ were annealed in either $^{18}O_2$ or $^{16}O_2$ gas under the same thermal conditions: 800 °C at 1 atmosphere for 168 h, followed by 700 °C at 0.2 bar for 24 h. The T_c values of the samples decreased from 92 K to 91 K on isotope substitution. Raman spectra showed a clear shift of the B_{1g} vibration modes, confirming that isotope substitution was nearly complete at 80% or more.

John F. Douglas*, Hideaki Iwasawa†‡, Zhe Sun*, Alexei V. Fedorov§, Motoyuki Ishikado¶, Tomohiko Saitoh†, Hiroshi Eisaki‡, Hiroshi Bando‡, Takeshi Iwase‡, Akihiro Ino||, Masashi Arita#, Kenya Shimada#, Hirofumi Namatame#, Masaki Taniguchi||#, Takahiko Masui, Setsuko Tajima**, Kazuhiro Fujita¶, Shin-ichi Uchida¶, Yoshihiro Aiura‡, Daniel S. Dessau***

*Department of Physics, University of Colorado, Boulder, Colorado 80309-0390, USA
†Department of Applied Physics, Tokyo University of Science, Tokyo 162-8601, Japan
‡National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki 305-8568, Japan
§Advanced Light Source, Lawrence Berkeley Lab, Berkeley, California 94720, USA.
¶Department of Physics, University of Tokyo, Tokyo 113-8656, Japan
||Graduate School of Science, Hiroshima University, Higashi-Hiroshima 739-8526, Japan
#Hiroshima Synchrotron Radiation Center, Hiroshima University, Higashi-Hiroshima 739-0046, Japan
**Department of Physics, Osaka University, 1-1 Machikaneyama, Toyonaka, Osaka 560-0043, Japan
e-mail: y.aiura@aist.go.jp

Received 29 September 2006; accepted 15 February 2007.

1. Gweon, G.-H. *et al. Nature* **430,** 187–190 (2004).
2. Franck, J. P. in *Physical Properties of High Temperature Superconductors IV* (ed. Ginsberg, D. M.) 189–293 (World Scientific, Singapore, 1994).
3. Lee, J. *et al. Nature* **442,** 546–550 (2006).
4. de Lozanne, A. *Nature* **442,** 522–523 (2006).

Competing financial interests: declared none.
doi:10.1038/nature05738

ARTICLES

A new eutriconodont mammal and evolutionary development in early mammals

Zhe-Xi Luo[1,2], Peiji Chen[3], Gang Li[3] & Meng Chen[2]

Detachment of the three tiny middle ear bones from the reptilian mandible is an important innovation of modern mammals. Here we describe a Mesozoic eutriconodont nested within crown mammals that clearly illustrates this transition: the middle ear bones are connected to the mandible via an ossified Meckel's cartilage. The connected ear and jaw structure is similar to the embryonic pattern in modern monotremes (egg-laying mammals) and placental mammals, but is a paedomorphic feature retained in the adult, unlike in monotreme and placental adults. This suggests that reversal to (or retention of) this premammalian ancestral condition is correlated with different developmental timing (heterochrony) in eutriconodonts. This new eutriconodont adds to the evidence of homoplasy of vertebral characters in the thoraco-lumbar transition and unfused lumbar ribs among early mammals. This is similar to the effect of homeobox gene patterning of vertebrae in modern mammals, making it plausible to extrapolate the effects of *Hox* gene patterning to account for homoplastic evolution of vertebral characters in early mammals.

Class Mammalia
Order Eutriconodonta (ref. 1)
Family Jeholodentidae (nov.)
***Yanoconodon allini* gen. et sp. nov.**

Etymology. *Yan* is for the Yan mountains in Northern Hebei Province; *conodon* is Latin for 'cuspate tooth', a common suffix for mammalian taxonomic names; *allini* refers to Edgar Allin's studies on mammalian ear evolution.

Holotype. Nanjing University - Paleontology Laboratory NJU-P 06001, preserved on a slab (NJU-P06001A, shown in Fig. 1) and counter-slab (NJU-P06001B, shown in the Supplementary Information) of laminated siltstones.

Locality and horizon. Yixian Formation at Daluozigou locality (41° 49′ N and 116° 31′ E) in Fengning County in Hebei Province, China. Equivalent beds in the western Liaoning Province were dated to 125 to 122 million years ago[2–4].

Diagnosis. Dental formula $I^2.C^1.P^2.M^3/I_2.C_1.P_2.M_3$, with plesiomorphic 'triconodont' molars having three laterally compressed main cusps in straight alignment. Among mammaliaforms with 'triconodont-like' postcanines, it differs from *Sinoconodon* and morganucodontans in its absence of a post-dentary trough[1,5,6]. Among eutriconodontans, it differs from all known 'amphilestids' and gobiconodontids in its absence of lower cingular cuspules e and f and the upper lingual cingulum[1,5–12]. *Yanoconodon* and *Jeholodens*[13] are identical in molar characteristics and clustered in one clade (Fig. 2)[14–16]. *Yanoconodon allini* differs from *Jeholodens jenkinsi* in having a triangular outline of scapula and several differences in thoracolumbar vertebrae (full diagnoses of *Y. allini* and the family Jeholodentidae are given in the Supplementary Information).

Description

The last molariform in the holotype of *Yanoconodon* has partially erupted (Fig. 3b), similar to the condition in *Jeholodens*[13]. However, the exposed interior of the lower and upper jaws shows no replacement at any tooth loci and its functional teeth (including the last tooth) are all permanent. The holotype is either a late-stage subadult or an adult with a delayed eruption of the last molar. It differs from gobiconodontids in the absence of replacements at the anterior molariform loci, although it shares this delayed eruption of the ultimate molariform with older gobiconodontid individuals[8–12].

Yanoconodon is preserved with the middle-ear bones (except the stapes), an ossified Meckel's cartilages, and the associated (although disarticulated) hyoid elements (Fig. 1; also Supplementary Information). The malleus has a manubrium that is similar to (although more robust than) that of the adult monotreme *Ornithorhynchus*[17–19], and the goniale element ('prearticular') is also present (Fig. 3). Similarly to modern mammals, the incus has a crus longum (stapedial process) and a crus breve (for basicranial articulation). The ectotympanic ring ('reflected lamina of angular') forms an arc of about 90°. The dorsal crus of the ectotympanic and the prearticular element of the malleus are fused to each other and both are connected anteriorly with the ossified Meckel's cartilage (Fig. 3b, g). The Meckel's cartilage has a mediolaterally compressed anterior (mandibular) limb, and a dorsoventrally compressed posterior (tympanic) limb. These two limbs are twisted relative to each other and curved at the mid-length of the cartilage (red arrow in Fig. 3). By comparison to *Yanoconodon*, we recognize that an ossified Meckel's cartilage is also preserved, although it is detached from the mandible in *Jeholodens*[13]. The Meckel's cartilage is similar to those preserved in several gobiconodontids (Fig. 3e, 3f)[9–12] but adds new information on its structural connection with the middle-ear bones, which are as yet unknown in gobiconodontids.

Limb and foot structure

Yanoconodon has a triangular outline of the scapula, and a gracile and slightly curved clavicle that lacks a rigid articulation to the interclavicle, it therefore has a mobile and 'therian-like' shoulder girdle (Fig. 1). However, the humerus resembles those of mammaliaforms or cynodonts[20–25] in a spindle-shaped (instead of spherical) head, a broad and shallow intertubercular groove, broad ect- and entepicondyles, and a proximo-distal torsion. The femur is similar to those of

[1]Carnegie Museum of Natural History, Pittsburgh, Pennsylvania 15213, USA. [2]Department of Earth Sciences, Nanjing University, Nanjing 210093, China. [3]State Key Laboratory of Palaeobiology and Stratigraphy, Nanjing Institute of Geology and Palaeontology, Chinese Academy of Sciences, Nanjing 210008, China.

cynodonts in that the femoral head is not spherical and has no neck; the greater trochanter has no vertical process. Gobiconodontids and jeholodentids lack the epiphyseal growth plates between the epiphyses (growth caps) and diaphysis (shaft) of long bones, a plesiomorphy of mammaliaforms and cynodonts[20–25]. But such growth plates are typical in multituberculates, extant monotremes, and therians[26–33]. Many modern therians have relatively rapid early skeletal growth, but terminate this growth upon the fusion of the epiphyses to the diaphysis. By contrast, eutriconodonts lacked a similar mechanism for terminating skeletal growth that would presumably slow down (although not stop completely) in the adult, because of the absence of the epiphyseal growth plates. The astragalus is oblong in outline without a distinctive neck or a pulley-like articulation for the upper ankle joint. The calcaneus has a broad peroneal shelf and a short tuber. *Yanoconodon* and *Jeholodens* lack a long list of apomorphies of limb and foot bones of multituberculates and therian mammals[26–35]. The fore- and hind-limbs both have sprawling posture (Fig. 1b). The metacarpals, metatarsals, and all phalanges are short and stout, which are primitive features also of cynodonts, mammaliaforms, gobiconodontids and *Jeholodens*[12,13,20–25], for either terrestrial or fossorial habits.

The vertebral column of *Yanoconodon* (NJU-P06001A and B) has seven cervical vertebrae with unfused ribs. Components of the atlas and axis are not co-ossified (Fig. 1). Owing to a gradational transition from thoracics to lumbars (Figs 1 and 5), designation of the 18 thoracic versus eight lumbar vertebrae is somewhat arbitrary, based on the criterion that lumbars tend to have larger and more robust centra and wider rib 'plates'. *Yanoconodon* is similar to gobiconodontids[12,36] but different from closely related *Jeholodens* in having these primitive features. Similar to many cynodonts and some mammaliaforms[20–25], the anterior seven lumbar vertebrae have mobile ribs, several with an expanded proximal portion of the rib as in (but less developed than) the plated lumbar ribs of many cynodonts and the docodont *Castorocauda*[22–25]. *Yanoconodon* and *Repenomamus*[36] have 26 thoracolumbar vertebral segments, an exceptional number if compared to the 22 thoracolumbars of *Jeholodens*, and 19 or 20 thoracolumbars in most Tertiary and modern mammals[37].

Implications for mammalian ear evolution

The rare preservation of Meckel's cartilage in association with the ectotympanic, malleus and incus in *Yanoconodon* provides the following new observation of the middle ear and its relationship to the mandible. The outline and proportion of the ectotympanic, malleus and incus of *Yanoconodon* are similar to their homologues in adult *Ornithorhynchus* (except for the gracile tips of the manubrium and ectotympanic in the latter) (Fig. 3c, 3h)[17–19]. *Yanoconodon* is far more derived than mammaliaforms (Fig. 3a)[5,6,22] in that its middle-ear bones are mediolaterally separated from the pterygoid part of the mandible (Fig. 3b, 3h), despite the plesiomorphic similarity in retaining the Meckel's connection to the mandible.

The homology of post-dentary elements of non-mammalian tetrapods with the mammalian middle-ear bones has long been established[38–44]. Long before the current fossil evidence was discovered, it was hypothesized that the migration of middle-ear bones from the

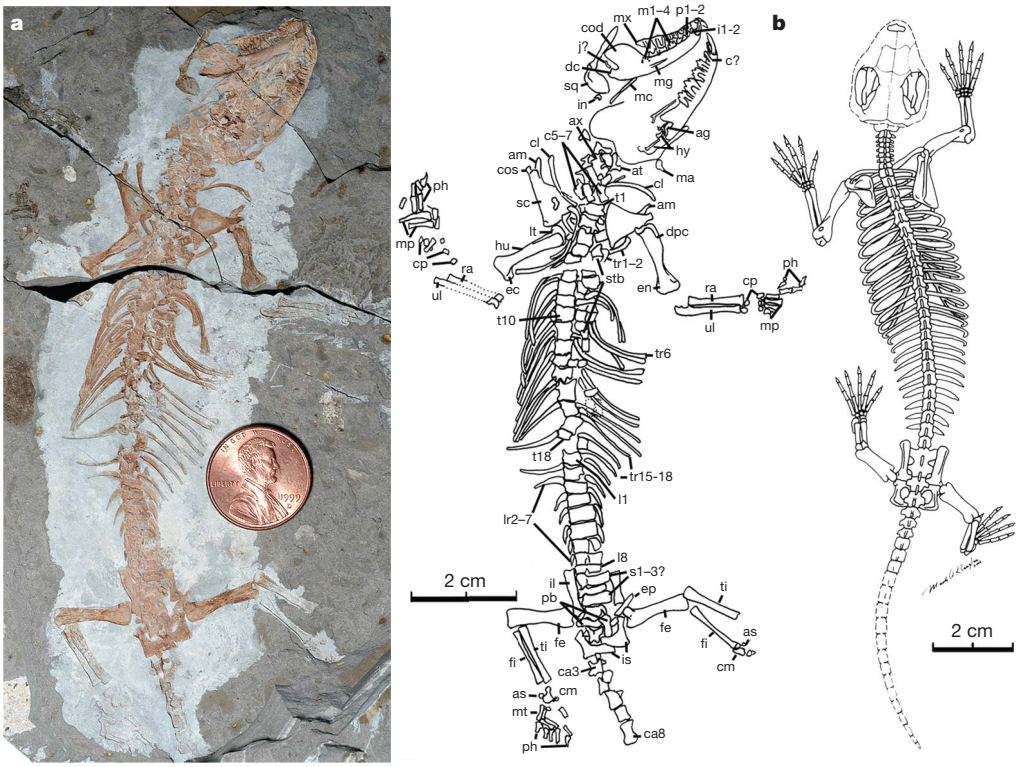


Figure 1 | New mammal *Yanoconodon allini*. **a**, Main part of the holotype (Nanjing University-Paleontology NJU-P06001A; the counterpart NJU-P06001B is illustrated in the Supplementary Information). **b**, Skeletal restoration (dorsal view). Abbreviations: ag, angular (ectotympanic ring); am, acromion (scapula); as, astragalus; at, atlas (cervical vertebra, c1); ax, axis (c2); c?, canine?; c5–7, cervical vertebrae 5–7; ca3 and ca8, caudal vertebrae 3 and 8 (distal caudals are missing); cl, clavicle; cm, calcaneum; cod, coronoid process (dentary); cos, coracoid process (scapula); cp, carpals; dc, dentary condyle; dpc, deltopectoral crest (humerus); ec, ectepicondyle (humerus); en, entepicondyle; ep, epipubis; fe, femur; fi, fibula; hu, humerus; hy, hyoid elements; i1–2, incisors 1 and 2; il, ilium; in, incus; is, ischium; j?, jugal?; l1 and l8, lumbar vertebrae 1 and 8; lr2–7, lumbar ribs 2–7; lt, lesser tubercle (humerus); m1–4, lower molars 1–4; ma, malleus; mc, Meckel's cartilage (ossified); mg, Meckel's groove (dentary); mp, metacarpals; mt, metatarsals; mx, maxillary; p1–2, premolars 1 and 2; pb, pubic; ph, phalanges; ra, radius; sc, scapula; sq, squamosal; stb, sternum and sternabrae; s1–3?, sacral vertebrae 1 and 2 (and possibly sacral vertebra 3?); t1, t10 and t18, thoracic vertebrae 1, 10 and 18; ti, tibia; tr1–2, tr6 and tr15–18, thoracic ribs 1, 2, 6 and 18; ti, tibia; ul, ulna.

mandible in cynodont–mammal evolution underwent an intermediate stage in which these bones would be anteriorly connected to the lower jaw, but suspended medial to (and free of) the mandibular pterygoid region (ref. 43 and Fig. 28.11 of ref. 44). *Yanoconodon* corroborates Allin's hypothetical model of an intermediate evolutionary stage. Newly discovered fossils of *Yanoconodon* (Fig. 1) and gobiconodontids[9–11], preserved in multiple specimens from two different fossil sites, add some crucial and previously unknown structural elements to this 'transitional' middle-ear model:

(1) The anterior (mandibular) limb and the posterior (tympanic) limb of Meckel's cartilage are twisted and curved relative to each other at the mid-length of the cartilage (red arrow in Fig. 3). (2) The mid-length twist and curvature of Meckel's cartilage made it feasible for the anterior limb of the cartilage to be nestled in the Meckel's groove on the mandible[9–11], while the tympanic limb of the cartilage and its associated ectotympanic and prearticular are separated mediolaterally from the pterygoid region of the mandible (Fig. 3f, g), as postulated by Allin[43]. The tympanic membrane suspended by the ectotympanic ring and malleus manubrium would be re-oriented more mediodorsally (see red arrows in b-1 and b-2 of Fig. 3b), similar to the ectotympanic's migratory path during ontogeny in the marsupial *Monodelphys*[41], leaving a space between the pterygoid region of the mandible and the tympanic membrane for access of the external auditory meatus and for pterygoid muscle. (3) Owing to the curvature of Meckel's cartilage, the basicranial articulation of the incus is nearly co-axial with the fulcrum for movement around the dentary-squamosal jaw hinge, so the jaw movement had little impact on the middle ear function. (4) The middle ear and the ossified Meckel's cartilage in late-stage subadult or adult of *Yanoconodon* (Fig. 3g) show a similar structural pattern to the embryos of monotreme (Fig. 3i) and placental mammals. The middle ear is anteriorly connected to a curved Meckel's cartilage, which is in turn connected to the mandible. In the meantime, the embryonic middle-ear bones are mediolaterally separated from the pterygoid region[17,18,39,40].

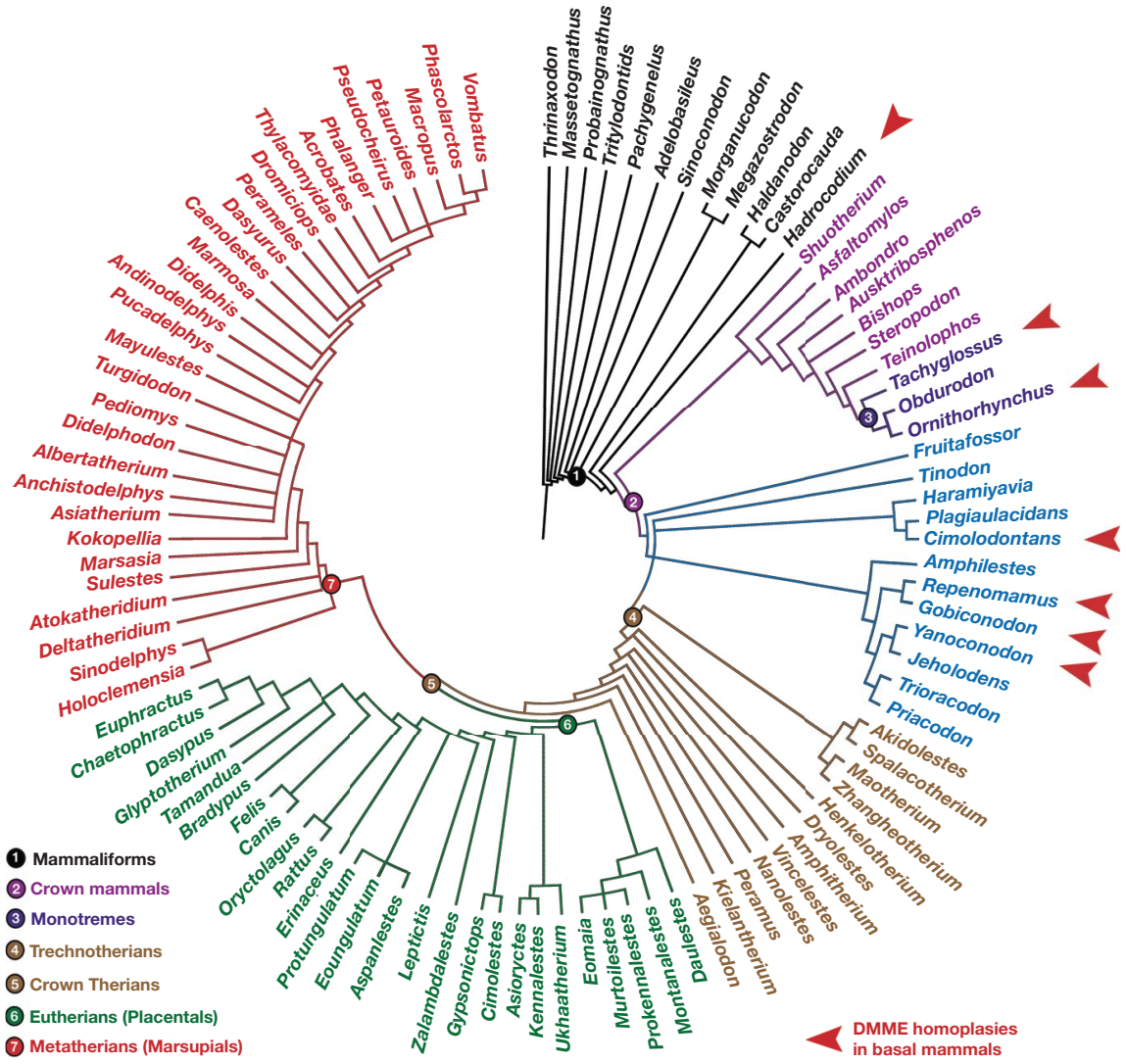


Figure 2 | Relationship of *Yanoconodon allini*. Analysis of all dental and skeletal characters of *Yanoconodon* places it as a close relative of *Jeholodens*, in the Jeholodentidae among eutriconodontans. This is based on the strict consensus of 218 equally parsimonious trees of PAUP (Phylogenetic Analysis Using Parsimony and Other Methods, version 4.0b) analysis of 436 characters (1,000 heuristic runs with unordered multi-state characters) that can be scored for 102 comparative taxa (97 mammaliaforms including 25 extant mammals, plus three cynodonts as outgroups). For each equally parsimonious tree, tree length = 2,188, consistency index = 0.375, retention index = 0.803. Characters are based on refs 1, 14–16, 22, 33 and this study. Graphic produced by M. Klingler.

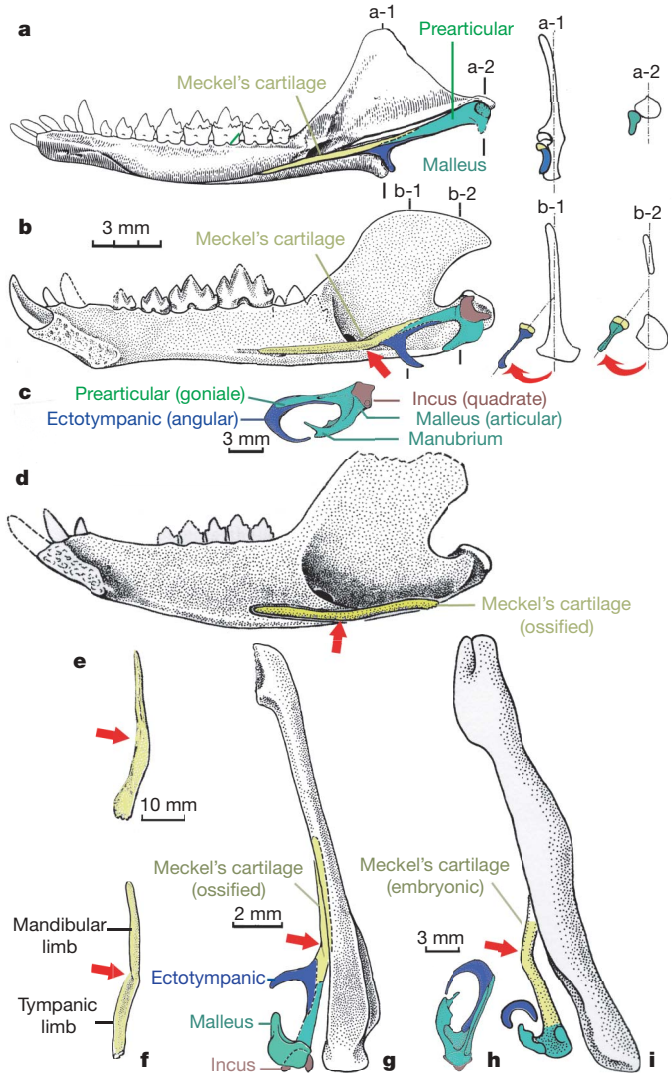


Figure 3 | Comparison of mandible and middle ear of *Yanoconodon*. **a**, Mammaliaform *Morganucodon* (medial view); a-1 and a-2 label schematic transverse sections at the levels of the malleus and the ectotympanic[5–7]. In *Morganucodon*, the middle ear maintains both an anterior connection to the mandible via the Meckel's cartilage, and a mediolateral contact to the mandible. **b**, Eutriconodont *Yanoconodon* (medial view, composite restoration of mandible and middle ear from NJU-P06001A and B). b-1 and b-2 label transverse sections at the levels of the malleus and the ectotympanic. The middle ear retains the anterior connection to the mandible via ossified Meckel's cartilage (yellow), but is mediolaterally separated from the posterior part of the mandible because of the twist and curvature of Meckel's cartilage (red arrows in **b**). **c**, the ectotympanic (blue), malleus (green) and incus (brown) of modern *Ornithorhynchus*[18]: the shape and proportion of the ear bones are similar in *Ornithorhynchus* and *Yanoconodon*. **d**, Eutriconodont *Repenomamus*[9–11]: ossified Meckel's cartilage connected anteriorly to the mandible (similar to *Yanoconodon*). **e**, Ossified Meckel's cartilage of *Repenomamus* (ventral view, isolated, redrawn from Fig. 1 of ref. 9). **f**, Ossified Meckel's cartilage of *Yanoconodon* (ventral view, isolated, composite restoration of both the left and the right elements). **g**, Middle ear of *Yanoconodon* (composite restoration, ventral view): the ectotympanic and malleus are connected anteriorly to the mandible via ossified Meckel's cartilage; but these are mediolaterally separated from the posterior part of mandible, facilitated by curvature of the Meckel's cartilage (yellow). **h**, Middle ear bones of adult *Ornithorhynchus* (ventral view) and similarity to those of *Yanoconodon*. **i**, Embryonic *Ornithorhynchus*[17]: the tympanic ring and the partially developed manubrium and goniale ('prearticular') of the malleus are anteriorly connected via Meckel's cartilage to the mandible, but separated mediolaterally from the posterior region of mandible, facilitated by the curved cartilage (red arrow). *Yanoconodon* retains the embryonic pattern of *Ornithorhynchus* owing to the timing change of earlier ossification of Meckel's cartilage, but otherwise its ectotympanic, malleus and incus are nearly the same as in adult *Ornithorhynchus*.

Overall, *Yanoconodon*'s middle ear shows a paedomorphic resemblance to the embryonic pattern of modern mammals. All that is necessary for adult *Yanoconodon* to retain this paedomorphic (albeit permanent) connection of the middle ear to the mandible is a relatively earlier timing in (premature) ossification of the Meckel's cartilage and its fusion with the ectotympanic (Fig. 3c–g).

Developmental heterochrony and evolutionary homoplasy

A definite mammalian middle ear (DMME)[43,44] is accomplished by two ontogenetic steps in extant monotremes and placentals: first, a mediolateral separation of the middle ear bones from the mandible in early embryonic stages (secondary mediolateral contact in marsupials due to inflected mandibular angle; see Fig. 12.10 of ref. 40), and second, a loss of anterior connection to the mandible owing to the reabsorption of Meckel's cartilage in the subadult. In *Yanoconodon* and gobiconodontids (and by inference also *Jeholodens*), the former step of mediolateral separation of the middle ear from the mandible has already occurred. But reabsorption of Meckel's cartilage did not happen, resulting in the retention in *Yanoconodon* of the middle-ear connection to the mandible, otherwise seen only in the early embryonic or fetal stage of extant mammals[17,18,39–42].

Yanoconodon and its eutriconodontan kin are nested within the crown Mammalia (Fig. 2) by the parsimony of all characters[14–16]. The absence of DMME in eutriconodonts, an in-group of crown Mammalia, is in sharp contrast to modern monotremes and therians that have DMME. This phylogeny requires one of the following two evolutionary scenarios: either (1) DMME was present in the common ancestor of monotremes, eutriconodonts and therians; but eutriconodonts re-evolved the middle ear attachment to mandible, or (2) DMME was absent in the common ancestor of monotremes, eutriconodonts and therians, and this is retained as a paedomorphosis in eutriconodonts; but DMME evolved in extant monotremes[44–47], and separately in therians[39]. Paedomorphosis, or retention of fetal or juvenile characteristics of ancestors and relatives through developmental heterochrony, is a common phenomenon in vertebrate evolution. The heterochronic ('premature') ossification of Meckel's cartilage in eutriconodonts is the immediate cause for this paedomorphic connection of middle ear and mandible, and is why there is an overall homoplastic distribution among therians (with DMME), eutriconodonts (without DMME), monotremes (with DMME) and pre-mammalian relatives (without DMME) (triangles in Fig. 2). The paedomorphic connection of the middle ear to mandible of eutriconodonts and mammaliaforms is consistent with their lack of the long-bone epiphyses for terminating skeletal growth, as seen in modern mammals.

Hox gene patterning of axial skeleton in extinct mammals

Modern mammals have a highly conserved pattern of vertebral identities: seven cervical, 13 to 14 thoracic, and five or six lumbar vertebrae (without separate lumbar ribs) for a combined 19 or 20 thoracolumbar vertebrae[37]. These regional vertebral identities are patterned by homeobox genes[37,48–50]. Homeotic changes in vertebral identities, such as shift of the thoracolumbar boundary or gradational transition, are now correlated with the loss and gain of *Hox* genes function in mice[48]. The distinctive boundary of thoracic versus lumbar regions and the absence of lumbar ribs are patterned by the *Hox10* paralogues in modern placental mammals (Fig. 4)[48]. The triple knockout of *Hox10* paralogues can alter the thoracic versus lumbar boundary, and triple knockout of *Hox11* can alter lumbar versus sacral vertebral identities. A loss of *Hox10* gene function can regenerate the lumbar ribs and a more gradational thoraco-lumbar transition in laboratory mice (Fig. 4)[48].

Yanoconodon and *Repenomamus* both have 26 thoracolumbar vertebrae, and these eutriconodontans represent a high variation for the highly conserved 19 to 20 thoracolumbar vertebrae of crown Mammalia. By contrast, *Jeholodens* lost lumbar ribs and achieved a

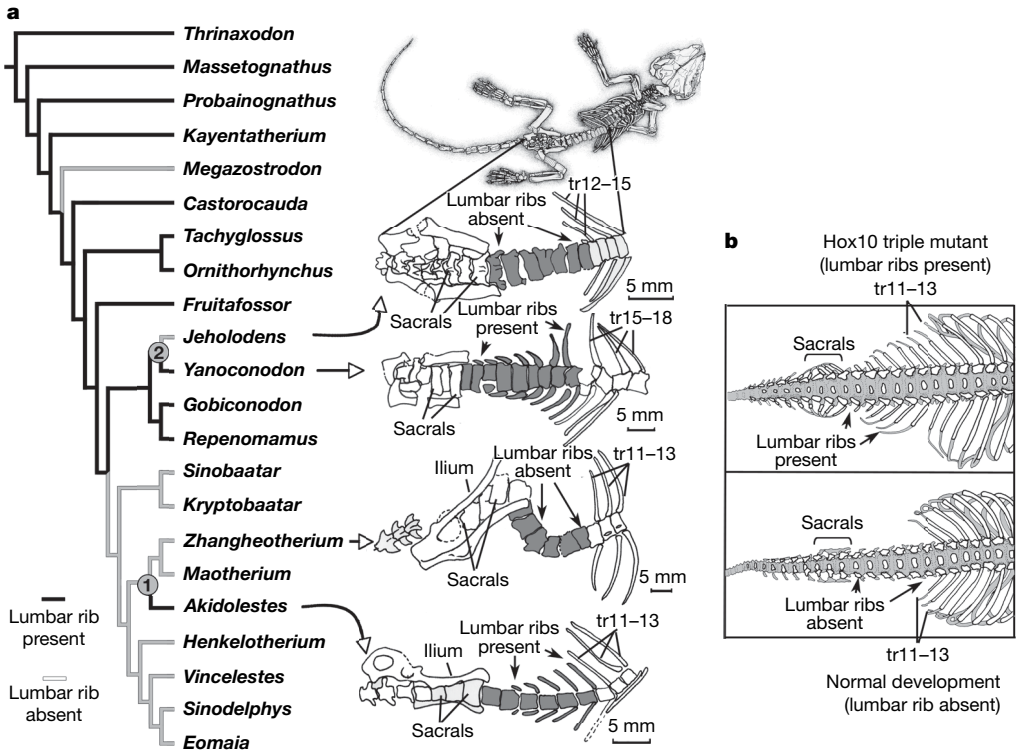


Figure 4 | Homoplastic evolution of lumbar ribs among Mesozoic mammals and patterning of vertebral and rib development by *Hox10* gene in extant mammals. a, Homoplastic distribution of lumbar ribs in Mesozoic mammal taxa preserved with vertebral column (topology from Fig. 2): lumbar ribs are present in gobiconodontids and *Yanoconodon* but absent in closely related *Jeholodens* (node marked 2), present in *Akidolestes* but absent in closely related *Zhangheotherium*[33] and the more inclusive theriiforms (node 1). **b**, Patterning of vertebral structure (development of lumbar ribs) in modern laboratory mice by *homeobox* genes (redrawn and flipped from Fig. 1 of ref. 48). A separate loss of lumbar ribs in *Jeholodens* among eutriconodontans is hypothesized to be correlated with an independent activation of *Hox10* patterning of thoracolumbar vertebrae (node 2). An isolated occurrence of lumbar ribs in *Akidolestes* among most spalacotheroids without lumbar ribs is hypothesized to be the effect of an independent loss of *Hox10* gene function. The loss or gain of *Hox* gene function to pattern the vertebral identities is a plausible mechanism for homoplasy of lumbar ribs in early mammals, and for variation of thoracolumbar vertebral counts among eutriconodontans. tr, numbered thoracic ribs.

distinctive boundary between the thoracic and lumbar regions (Fig. 4: node 2); its 15 thoracic and 7 lumbar vertebral counts are more comparable to modern mammals. These derived lumbar features of *Jeholodens* evolved from within eutriconodontans, convergent to those in theriiform mammals (node 1 in Fig. 4), and may be hypothesized as an independent activation of *Hox10* gene patterning[48].

Homoplasy of mobile lumbar ribs also occurs in symmetrodonts (Fig. 4)[33]. The presence of primitive lumbar ribs in *Akidolestes* is best hypothesized as an atavistic reversal because its related spalacotheroids show the derived therian condition without lumbar ribs[29,32,33]. Given that the regeneration of lumbar ribs in mutant mice is correlated with loss of function of *Hox10* paralogues (Fig. 4b), it is plausible that a similar mutation (that is, triple knockout of *Hox10* paralogues or a similar developmental event) was correlated with the 'reappearance' of lumbar ribs in *Akidolestes*, whose immediate relatives in successive ranks of theriiform mammals (node 1 in Fig. 4) all lack lumbar ribs.

Shifts in the thoracolumbar boundary and homoplasy of their vertebral identities are primarily correlated with mutation in *Hox* genes for their patterning[37,48–50]. This provides a plausible mechanism for the evolutionary patterns in lumbar ribs and the thoracolumbar transition in Mesozoic mammals (Fig. 4). This developmental mechanism for evolutionary homoplasy of thoracolumbar structure is mutually compatible with one of the two functional interpretations: either the lumbar ribs in *Akidolestes* had a function similar to that in primitive mammals and cynodonts with similar ribs[23–25], or the separate loss of lumbar ribs in *Jeholodens* was an adaptation convergent to most theriiform mammals. Reciprocally, homoplastic thoraco-lumbar characters within eutriconodontans and symmetrodontans[29,32,33] provide two cases for extrapolating the *Hox* gene patterning of laboratory mice to early mammal phylogeny on a grand evolutionary scale.

Received 4 August 2006; accepted 29 January 2007.

1. Kielan-Jaworowska, Z. *et al. Mammals from the Age of Dinosaurs—Origins, Evolution, and Structure* (Columbia Univ. Press, New York, 2004).
2. Ji, Q. *et al. Mesozoic Jehol Biota of Western Liaoning, China* (Geol. Publ. House, Beijing, 2004).
3. Zhou, Z.-H. *et al.* An exceptionally preserved Lower Cretaceous ecosystem. *Nature* **421,** 807–814 (2003).
4. Zhang, F. *et al.* Description of a new enantiornithine bird from the Early Cretaceous of Hebei, northern China. *Can. J. Earth Sci.* **41,** 1097–1107 (2004).
5. Kermack, K. A. *et al.* The lower jaw of *Morganucodon. Zool. J. Linn. Soc. (Lond.)* **53,** 87–175 (1973).
6. Crompton, A. W. & Luo, Z.-X. in *Mammal Phylogeny* Vol. 1 (eds Szalay, F. S. *et al.*) 30–44 (Springer, New York, 1993).
7. Luo, Z.-X. *et al.* A new mammaliaform from the Early Jurassic of China and evolution of mammalian characteristics. *Science* **292,** 1535–1540 (2001).
8. Jenkins, F. A. Jr & Schaff, C. R. The Early Cretaceous mammal *Gobiconodon* (Mammalia, Triconodonta) from the Cloverly Formation in Montana. *J. Vert. Paleontol.* **8,** 1–24 (1988).
9. Wang, Y.-Q. *et al.* An ossified Meckel's cartilage in two Cretaceous mammals and origin of the mammalian middle ear. *Science* **294,** 357–361 (2001).
10. Li, C.-K. *et al.* A new species of *Gobiconodon* (Triconodonta, Mammalia) and its implication for the age of Jehol Biota. *Chin. Sci. Bull.* (English edn) **48,** 1129–1134 (2003).
11. Meng, J. *et al.* The ossified Meckel's cartilage and internal groove in Mesozoic mammaliaforms: implications to origin of the definitive mammalian middle ear. *Zool. J. Linn. Soc. (Lond.)* **138,** 431–448 (2003).
12. Hu, Y.-M. *et al.* Large Mesozoic mammals fed on young dinosaurs. *Nature* **433,** 149–153 (2005).
13. Ji, Q. *et al.* A Chinese triconodont mammal and mosaic evolution of the mammalian skeleton. *Nature* **398,** 326–330 (1999).
14. Luo, Z.-X. *et al.* Dual origin of tribosphenic mammals. *Nature* **409,** 53–57 (2001).
15. Luo, Z.-X. *et al.* In quest for a phylogeny of Mesozoic mammals. *Acta Palaeont. Polonica* **47,** 1–78 (2002).

16. Luo, Z.-X. & Wible, J. R. A new Late Jurassic digging mammal and early mammalian diversification. *Science* **308,** 103–107 (2005).
17. Zeller, U. Die Entwicklung und Morphologie des Schädels von *Ornithorhynchus anatinus* (Mammalia: Prototheria: Monotremata). *Abhandl. Senckenberg. Natur. Gesell.* **545,** 1–188 (1989).
18. Zeller, U. in *Mammal Phylogeny* Vol. 1 (eds Szalay, F. S. *et al.*) 95–107 (Springer, New York, 1993).
19. Fleischer, G. Studien am Skelett des Gehörorgans der Säugetiere, einschliesslich des Menschen. *Säugetierk. Mitteil.* **21,** 131–239 (1973).
20. Jenkins, F. A. Jr & Parrington, F. R. The postcranial skeletons of the Triassic mammals *Eozostrodon, Megazostrodon* and *Erythrotherium. Phil. Trans. R. Soc. Lond.* **273,** 387–431 (1976).
21. Martin, T. Postcranial anatomy of *Haldanodon exspectatus* (Mammalia, Docodonta) from the Late Jurasssic (Kimmeridgian) of Portugal and its bearing for mammalian evolution. *Zool. J. Linn. Soc. (Lond.)* **145,** 219–248 (2005).
22. Ji, Q. *et al.* A swimming mammaliaform from the Middle Jurassic and ecomorphological diversification of early mammals. *Science* **311,** 1123–1127 (2006).
23. Jenkins, F. A. Jr. The postcranial skeleton of African cynodonts. *Peabody Mus. Nat. Hist. Bull.* **36,** 1–216 (1971).
24. Jenkins, F. A. Jr. The Chañares (Argentina) Triassic reptile fauna VII. The postcranial skeleton of the traversodontid *Massetognathus pascuali* (Therapsida, Cynotondia). *Breviora* **352,** 1–28 (1970).
25. Sues, H.-D. & Jenkins, F. A. Jr. in *Amniote Paleobiology: Perspectives on the Evolution of Mammals, Birds, and Reptiles* (eds Carrano, M. T. *et al.*) 114–152 (Univ. Chicago Press, Chicago, 2006).
26. Krebs, B. 1991. Das Skelett von *Henkelotherium guimarotae* gen. et sp. nov. (Eupantotheria, Mammalia) aus dem Oberen Jura von Portugal. *Berliner Geowisch. Abhandl.* **A133,** 1–110 (1991).
27. Rougier, G. W. Vincelestes neuquenianus *Bonaparte (Mammalia, Theria), un Primitivo Mamífero del Cretácico Inferior de la Cuenca Neuquina.* PhD dissertation (Univ. Nacional Buenos Aires, 1993).
28. Szalay, F. S. *Evolutionary History of the Marsupials and an Analysis of Osteological Characters* (Cambridge Univ. Press, Cambridge, 1994).
29. Hu, Y.-M. *et al.* A new symmetrodont mammal from China and its implications for mammalian evolution. *Nature* **390,** 137–142 (1997).
30. Ji, Q. *et al.* The earliest-known eutherian mammal. *Nature* **416,** 816–822 (2002).
31. Luo, Z.-X. *et al.* An Early Cretaceous tribosphenic mammal and metatherian evolution. *Science* **302,** 1934–1940 (2003).
32. Luo, Z.-X. & Ji, Q. New study on dental and skeletal features of the Cretaceous mammal *Zhangheotherium. J. Mammal. Evol.* **12,** 337–357 (2005).
33. Li, G. & Luo, Z.-X. A Cretaceous symmetrodont therian with some monotreme-like postcranial features. *Nature* **439,** 195–200 (2006).
34. Krause, D. W. & Jenkins, F. A. Jr. The postcranial skeleton of North American multituberculates. *Bull. Mus. Comp. Zool.* **150,** 199–246 (1983).
35. Kielan-Jaworowska, Z. & Gambaryan, P. P. Postcranial anatomy and habits of Asian multituberculate mammals. *Fossils Strata* **36,** 1–92 (1994).
36. Hu, Y.-M. The postcranium of *Repenomamus* and its implications for evolution of mammalian skeletal characters. *J. Vert. Paleontol.* **22** (3-Suppl.), 67A–68A (2002).
37. Narita, Y. & Kuratani, S. Evolution of vertebral formulae in mammals: a perspective on developmental constraints. *J. Exp. Zool.* **304B,** 91–106 (2005).
38. Gaupp, E. Die Reichertsche Theorie (Hammer-, Amboss- und Kieferfrage). *Archiv Anatomie Entwick.* **1912,** 1–426 (1913).
39. Maier, W. Phylogeny and ontogeny of mammalian middle ear structures. *Nether. J. Zool.* **40,** 55–75 (1990).
40. Maier, W. in *Mammal Phylogeny (Volume 1)* (eds Szalay, F. S. *et al.*) 165–181 (Springer, New York, 1993).
41. Rowe, T. B. Coevolution of the mammalian middle ear and neocortex. *Science* **273,** 651–654 (1996).
42. Sánchez-Villagra, M. R. *et al.* Ontogenetic and phylogenetic transformations of the ear ossicles in marsupial mammals. *J. Morphol.* **251,** 219–238 (2002).
43. Allin, E. F. Evolution of the mammalian middle ear. *J. Morphol.* **147,** 403–438 (1975).
44. Allin, E. F. & Hopson, J. A. in *The Evolutionary Biology of Hearing* (eds Webster, D. B. *et al.*) 587–614 (Springer, New York, 1992).
45. Rich, T. H. *et al.* Independent origins of middle ear bones in monotremes and therians. *Science* **307,** 910–914 (2005).
46. Martin, T. & Luo, Z.-X. Paleontology: homoplasy in the mammalian ear. *Science* **307,** 861–862 (2005).
47. Bever, G. *et al.* Comment on ''Independent origins of middle ear bones in monotremes and therians '' (I-II). *Science* **309,** 1492a–1492b (2005).
48. Wellik, D. M. & Capecchi, M. R. *Hox10* and *Hox11* genes are required to globally pattern the mammalian skeleton. *Science* **301,** 363–367 (2003).
49. Burke, A. C. & Nowicki, J. L. Hox genes and axial specification in vertebrates. *Am. Zool.* **41,** 687–697 (2001).
50. Galis, F. Why do almost all mammals have seven cervical vertebrae? Developmental constraints, Hox genes, and cancer. *J. Exp. Zool.* **285,** 19–26 (1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Tabrum for preparing this fossil; X.-N. Yang, Y.-K. Shi, J.-R. Liu, Q. Yang, J.-G. Sha, H.-C. Zhang for support; Q. Ji and J. Wible for access to comparative collections; R. Cifelli, Z. Kielan-Jaworowska, T. Martin, T. Rowe, J. Wible and G. Wilson for discussions; M. Dawson and J. Wible for improving the manuscript; and M. Klingler for assistance with the figures. This work was supported by the National Natural Science Foundation of China (P.C., G.L. and Z.-X.L.), the National Science Foundation and National Geographic Society (Z.-X.L.), the Ministry of Science and Technology of China (the 973 Project under C.-S. Wang) and the State Key Laboratory of Palaeobiology and Stratigraphy of NIGPAS (G. L.).

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to Z.-X.L. (luoz@carnegiemnh.org).

LETTERS

A collisional family of icy objects in the Kuiper belt

Michael E. Brown[1], Kristina M. Barkume[1], Darin Ragozzine[1] & Emily L. Schaller[1]

The small bodies in the Solar System are thought to have been highly affected by collisions and erosion. In the asteroid belt, direct evidence of the effects of large collisions can be seen in the existence of separate families of asteroids—a family consists of many asteroids with similar orbits and, frequently, similar surface properties, with each family being the remnant of a single catastrophic impact[1]. In the region beyond Neptune, in contrast, no collisionally created families have hitherto been found[2]. The third largest known Kuiper belt object, 2003 EL_{61}, however, is thought to have experienced a giant impact that created its multiple satellite system, stripped away much of an overlying ice mantle, and left it with a rapid rotation[3–5]. Here we report the discovery of a family of Kuiper belt objects with surface properties and orbits that are nearly identical to those of 2003 EL_{61}. This family appears to be fragments of the ejected ice mantle of 2003 EL_{61}.

Near-infrared reflectance spectroscopy has shown a diversity of surface compositions on Kuiper belt objects (KBOs), ranging from surfaces dominated by methane absorptions, to those with water-ice absorptions, to those with no discernible infrared spectral features[6]. Currently, the processes that create these diverse surfaces are not well understood. In an effort to better quantify the spectral signatures seen on KBOs, we obtained near-infrared reflectance surface spectra of 30 KBOs at the W.M. Keck observatory (see Supplementary Information for all spectra). Our survey confirms the three broad classes of KBO surface types: the largest KBOs—Pluto, Eris and 2005 FY_9—have spectra dominated by methane-ice absorption bands[7–9], while the remaining objects are either spectrally dominated by absorption features due to water ice at 1.5 and 2.0 μm wavelength or are featureless in the infrared to the level of noise.

Examination of the non-methane objects reveals that although most show moderate or no water-ice absorption, six KBOs (namely the extremely large KBO 2003 EL_{61}, plus the much smaller 1995 SM_{55}, 1996 TO_{66}, 2002 TX_{300}, 2003 OP_{32} and 2005 RR_{43}), and also the brightest satellite of 2003 EL_{61} (ref. 4), show extremely deep absorption features characteristic of water ice. In addition, the measured colours of the surfaces of these KBOs are exclusively neutral, compared to the wide range of optical colours seen in the other KBOs[10]. An examination of the depth of the water-ice absorption as a function of colour (Fig. 1, see Supplementary Discussion for details) shows that these objects form a unique group with surface characteristics different from the remaining objects. When these iciest KBOs are excluded, no further correlation is seen between water absorption depth and colour. Using the Kuiper variant of the Kolmogorov–Smirnov test, we calculate that the likelihood that the KBOs with deep water-ice absorptions were selected from the same colour distribution as the remaining population is less than 1.2%.

As well as their surface characteristics suggesting similarities between the objects, the objects themselves are also clustered within an exceedingly small dynamical region of the Kuiper belt (Fig. 2). To more accurately examine the dynamical relationships between the objects, we determined their proper orbital elements by taking 50-Myr averages of their osculating orbital elements (see Supplementary Discussion). The proper elements differ by only a few per cent from each other: the semimajor axes (a) of the six objects with deep water-ice absorptions have a spread of 2.15 AU ($\Delta a/a \approx 0.05$), the eccentricities differ by 0.08, and the inclinations differ by 1.4° (0.02 rad). Four of these objects are even more tightly clustered, with differences of only 0.1 AU in semimajor axis, 0.03 in eccentricity, and 1.0° in inclination. These six objects have the smallest dispersion of any sample of six within our survey. The probability of randomly selecting the single most clustered set of six out of a sample of 32 is only 1×10^{-6}.

The similar surface characteristics and extremely small dynamical dispersion of these objects are naturally explained if all six KBOs were derived from a single disruptive collision. The largest of this group of KBOs, 2003 EL_{61}, has a mass that is probably of the order of 100 times more than the mass of the other five objects combined (assuming that the objects all have a similar albedo, which their similar surface characteristics suggest is reasonable, and that all of the objects have the same density, which must be correct within a factor of three). 2003 EL_{61} is thus the most plausible candidate for the remnant of the progenitor of this collisional family. Moreover, the fast rotation, multiple satellite system, and high density of 2003 EL_{61} have previously been argued to be due to a giant impact that ejected a fraction of the original icy mantle and left two satellites behind[3,4]. The discovery of a family of objects with very similar orbits to 2003 EL_{61} and with very similar surface properties to 2003 EL_{61} and its satellite

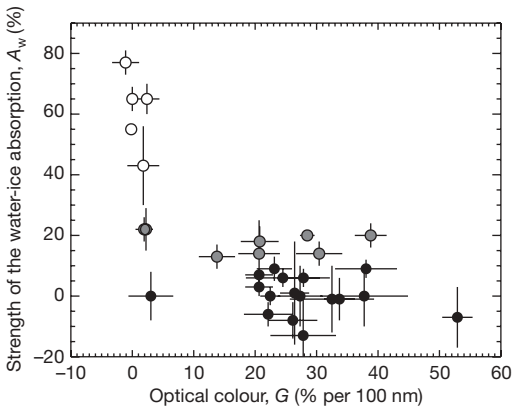


Figure 1 | The strength of the infrared water-ice absorption feature versus optical colour for all KBOs in our survey and those with published spectra. See Supplementary Information for the complete list of objects. The strength of the water-ice absorption feature, A_w, is defined as the fractional absorption at 2.0 μm compared to the reflectance at 1.8 μm. The optical colour, G, is defined as the fractional change in reflectance per 100 nm wavelength change between 0.5 and 0.8 μm. Six KBOs (open circles) appear clustered with large water-ice absorptions and neutral colours. The black dots indicate objects with absorption depth, A_w, less than 10%; the grey dots represent objects with A_w less than 30%. The white dots represent objects with A_w greater than 30%. Error bars are 1σ.

[1]Division of Geological and Planetary Sciences, California Institute of Technology, Mail Code 150-21, 1200 E. California Blvd, Pasadena, California 91125, USA.

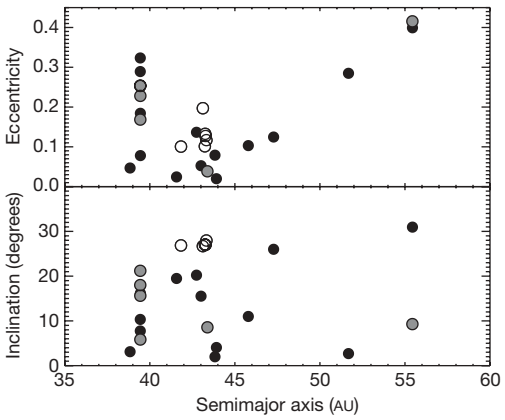


Figure 2 | The proper eccentricity and inclination of KBOs in our sample as a function of semimajor axis. The distributions in eccentricity and inclination of our population reflect the general distribution seen in the Kuiper belt. The black dots indicate objects with absorption depth, A_w, less than 10%; the grey dots represent objects with A_w less than 30%. The white dots represent objects with A_w greater than 30%. Four scattered objects with semimajor axes between 95 and 103 AU have been excluded so that the general distribution is clearly illustrated. The objects with the strongest water-ice absorption have clustered orbital parameters.

strongly argues that these family members are the dispersed remnants of the mantle of the proto-2003 EL_{61}.

To examine more closely whether the observed orbits could indeed result from collisional disruption, we construct a simple model of the aftermath of a dispersive collision by calculating the new orbital elements of test particles ejected from the dispersion of a proto-2003 El_{61}. We assume that the collision occurred as the proto-2003 EL_{61} crossed the ecliptic, where KBO number densities are the highest, and that, of the two ecliptic crossings in an inclined orbit, the impact occurred at the crossing at the higher-density portion of the Kuiper belt. The observed spread in orbital elements can be explained with ejection velocities of approximately 400 m s^{-1} (see Fig. 3). However, such velocities suggest that fragments should be strewn throughout a wide swath of the Kuiper belt, rather than confined to the small dynamical region observed. Although it is clear that the high-inclination population of the Kuiper belt is relatively unexplored and that many more fragments are likely to be eventually found, it appears improbable that the only fragments known would be so tightly clustered if such a large velocity dispersion occurred. In contrast, the orbital elements of all but one of the fragments can be explained by assuming that the centre of the tight cluster of four objects defines the centre of the collision, and that the collision dispersed objects with a velocity of only ~140 m s^{-1} above the primary escape velocity.

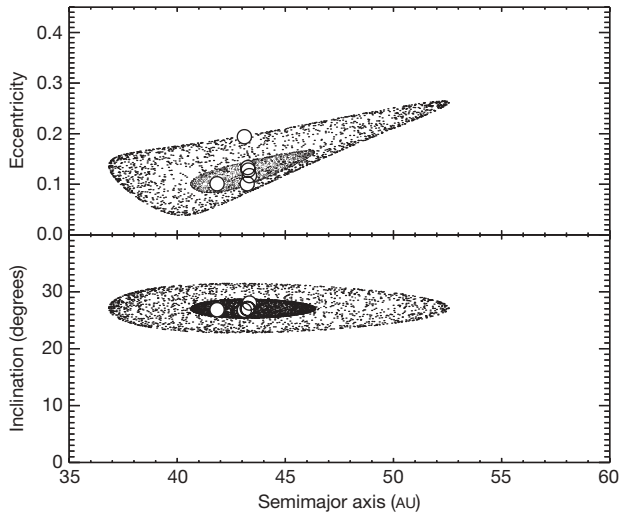


Figure 3 | Simulations of the dispersion in orbital elements expected after the ejection of fragments in a giant impact. The open circles give the current proper orbital elements of the fragments. The widely dispersed small dots show orbital elements expected from a dispersive velocity of 400 m s^{-1} and a dispersion centred on the average position of the fragments. Although this velocity can explain each of the known fragments, it also predicts that fragments should be strewn throughout a much larger region of the Kuiper belt. The more tightly concentrated dots show orbital elements expected from a dispersive velocity of 140 m s^{-1} and a dispersion centred on the average position of the central four KBOs. Five of the six fragments can be explained by this much smaller velocity. 2003 EL_{61}, the only object not fitted by this smaller velocity, is in a mean-motion resonance with Neptune that is capable of raising its eccentricity to its current value.

In this case, however, the single object that does not fit within the small velocity dispersion is 2003 EL_{61}, which still requires a velocity of ~400 m s^{-1} to explain its slight displacement in eccentricity from the remainder of the family. Long-term integration of the orbit of 2003 EL_{61} shows, however, that this object (and only this object) has large excursions in eccentricity over time owing to chaotic diffusion within the 12:7 mean-motion resonance with Neptune[11,12]. To explore this possibility, we integrated the orbits of 32 test particles with proper orbital elements within 0.3 AU in semimajor axis, 0.025 in eccentricity, and 0.5° in inclination from the centre of the cluster. We find that 5 of the 32 test particles diffuse out of the small initial region of the family and into higher-eccentricity orbits. All five of the test particles that diffused out appeared to be caught in the same 12:7 resonance that 2003 EL_{61} occupies. We thus suggest that 2003 EL_{61} was initially part of the extremely tightly clustered group of objects and that the initial collision placed it within the nearby 12:7 resonance, which subsequently raised the eccentricity to its current value.

Although the spectral and dynamical similarities of the objects argue for a common origin, and the rapid rotation, high density, and multiple satellite system of 2003 EL_{61} argue for a collisional history, some aspects of the system differ from general expectations. In general, models of the aftermath of collisions suggest that highly energetic impacts can either disrupt and disperse the primary or lead to the creation of a disk or satellite[13–17]. Simultaneous creation of both dispersed fragments and multiple satellites has not been seen. In addition, the 140 m s^{-1} dispersion between the objects is smaller than expected; simulations of disruptive impacts suggest that fragments are ejected with moderate fractions of the parent escape velocity[14,18], which would imply a dispersion a factor of three or more higher. Detailed simulations of collisions, however, show that a variety of outcomes are possible, depending on impact energies, impact angles, and compositions and initial spins of both bodies[17,19]. We can make a first-order estimate of the impactors involved by making plausible assumptions about the initial bodies. If the proto-2003 EL_{61} had a density of 2.0 g cm^{-3} (similar to that of Pluto, Triton, Eris and Charon, the only other Kuiper-belt-derived bodies of comparable size with known density), and the collision removed sufficient water ice to change the density to the current value of 2.6 g cm^{-3}, then the proto-2003 EL_{61} was a body of radius ~830 km, and ~20% of the initial mass was removed in the collision. At typical ecliptic encounter velocities of ~3 km s^{-1}, models suggest that a collision with an object 60% of the radius of the proto-2003 EL_{61} would cause such a moderate level mass removal[14]. For lower-velocity impacts, which may be required to explain the low velocity dispersion of the fragments, even larger impactors are required.

Finally, we consider the connection between a dispersive impact and the deep water-ice absorption. Although it is tempting to suggest that the deep water-ice absorption is a function of a relatively short surface exposure age for these objects, implying a relatively recent collision, little is understood about causes of spectral variations in KBOs. We note, however, that the deep water-ice absorptions are consistent with a relatively pure water-ice composition, as might be

expected from fragments that are the remnants of a dispersed icy mantle. In the end, however, we use the deep water-ice absorptions only as tracers of a common origin rather than as indications of a specific process related to an impact. It is possible that the currently recognized fragments are all from a similar region of the parent body, and that additional fragments with a variety of spectral signatures may some day be found.

Previous attempts to identify families of collisional fragments within the Kuiper belt using purely dynamical arguments, such as those usually used to identify asteroid belt families, have given generally ambiguous results[20]. The dual spectroscopic and dynamical approach used here, combined with the earlier suggestion that the largest member of this family independently shows signs of having experienced a giant impact, provides overwhelming circumstantial evidence for the identification of this collisional family.

Received 25 October 2006; accepted 19 January 2007.

1. Bendjoya, Ph. & Zappala, V. Asteroid family identification, in *Asteroids III* (eds Bottke, W., Cellino, A., Paolicchi, P. & Binzel, R. P.) 613–618 (Univ. Arizona Press, Tucson, 2002).
2. Farinella, P., Davis, D. R. & Stern, S. A. Formation and collisional evolution of the Edgeworth-Kuiper belt, in *Protostars and Planets IV* (eds Mannin, V., Boss, A. P. & Russell, S. S.) 1255–1268 (Univ. Arizona Press, Tucson, 2000).
3. Brown, M. E. *et al.* Satellites of the largest Kuiper belt objects. *Astrophys. J.* **639,** L43–L46 (2006).
4. Barkume, K. M., Brown, M. E. & Schaller, E. L. Water ice on the satellite of Kuiper belt object 2003 EL61. *Astrophys. J.* **640,** L87–L89 (2006).
5. Rabinowitz, D. L. *et al.* Photometric observations constraining the size, shape and albedo of 2003 EL61, a rapidly rotating Pluto-sized object in the Kuiper Belt. *Astrophys. J.* **639,** 1238–1251 (2006).
6. Barucci, M. A., Merlin, F., Dotto, E., Doressoundiram, A. & de Bergh, C. TNO surface ices: Observations of the TNO 55638 (2003 VE95) and analysis of the population's spectral properties. *Astron. Astrophys.* **455,** 725–730 (2006).
7. Brown, M. E., Trujillo, C. A. & Rabinowitz, D. L. Discovery of a planetary-sized object in the scattered Kuiper belt. *Astrophys. J.* **635,** L97–L99 (2005).
8. Licandro, J. *et al.* The methane ice rich surface of large TNO 2005 FY9: a Pluto-twin in the trans-neptunian belt? *Astron. Astrophys.* **445,** L35–L38 (2006).
9. Brown, M. E. *et al.* Methane and ethane on the bright Kuiper belt object 2005 FY9. *Astron. J.* (in the press).
10. Hainaut, O. R. & Delsanti, A. C. Colors of minor bodies in the outer Solar System. A statistical analysis. *Astron. Astrophys.* **389,** 641–644 (2002).
11. Nesvorny, D. & Roig, F. Mean motion resonances in the transneptunian region part II: The 1:2, 3:4, and weaker resonances. *Icarus* **150,** 104–123 (2001).
12. Kuchner, M. J., Brown, M. E. & Holman, M. Long-term dynamics and the orbital inclinations of the classical Kuiper belt objects. *Astron. J.* **124,** 1221–1230 (2002).
13. Melosh, H. J. & Ryan, E. V. Asteroids: Shattered but not dispersed. *Icarus* **129,** 562–564 (1997).
14. Benz, W. & Asphaug, E. Catastrophic disruptions revisited. *Icarus* **142,** 5–20 (1999).
15. Agnor, C. & Asphaug, E. Accretion efficiency during planetary collisions. *Astrophys. J.* **613,** L157–L160 (2004).
16. Canup, R. M. Simulations of a late lunar-forming impact. *Icarus* **168,** 433–456 (2004).
17. Canup, R. M. A giant impact origin of Pluto-Charon. *Science* **307,** 546–550 (2005).
18. Nesvorny, D. *et al.* Karin cluster formation by asteroid impact. *Icarus* **183,** 296–311 (2006).
19. Asphaug, E., Agnor, C. B. & Williams, Q. Hit-and-run planetary collisions. *Nature* **439,** 155–160 (2006).
20. Chiang, E. I. A collisional family in the classical Kuiper belt. *Astrophys. J.* **573,** L65–L68 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank R. Sari, B. McKinnon, K. Noll, T. Ahrens and A. Morbidelli for their suggestions and comments on this work. This research is supported by a grant to M.E.B. from NASA Planetary Astronomy.

Author Contributions M.E.B. was lead author of this Letter, and K.M.B., D.R. and E.L.S. contributed equally.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.E.B. (mbrown@caltech.edu).

LETTERS

Quantum jumps of light recording the birth and death of a photon in a cavity

Sébastien Gleyzes[1], Stefan Kuhr[1]†, Christine Guerlin[1], Julien Bernu[1], Samuel Deléglise[1], Ulrich Busk Hoff[1], Michel Brune[1], Jean-Michel Raimond[1] & Serge Haroche[1,2]

A microscopic quantum system under continuous observation exhibits at random times sudden jumps between its states. The detection of this quantum feature requires a quantum non-demolition (QND) measurement[1–3] repeated many times during the system's evolution. Whereas quantum jumps of trapped massive particles (electrons, ions or molecules[4–8]) have been observed, this has proved more challenging for light quanta. Standard photodetectors absorb light and are thus unable to detect the same photon twice. It is therefore necessary to use a transparent counter that can 'see' photons without destroying them[3]. Moreover, the light needs to be stored for durations much longer than the QND detection time. Here we report an experiment in which we fulfil these challenging conditions and observe quantum jumps in the photon number. Microwave photons are stored in a superconducting cavity for times up to half a second, and are repeatedly probed by a stream of non-absorbing atoms. An atom interferometer measures the atomic dipole phase shift induced by the non-resonant cavity field, so that the final atom state reveals directly the presence of a single photon in the cavity. Sequences of hundreds of atoms, highly correlated in the same state, are interrupted by sudden state switchings. These telegraphic signals record the birth, life and death of individual photons. Applying a similar QND procedure to mesoscopic fields with tens of photons should open new perspectives for the exploration of the quantum-to-classical boundary[9,10].

A QND detection[1–3] realizes an ideal projective measurement that leaves the system in an eigenstate of the measured observable. It can therefore be repeated many times, leading to the same result until the system jumps into another eigenstate under the effect of an external perturbation. For a single trapped ion, laser-induced fluorescence provides an efficient measurement of the ion's internal state[5–7]. The ion scatters many photons while evolving on a transition between a ground sublevel and an excited one. This fluorescence stops and reappears abruptly when the ion jumps in and out of a third, metastable level, decoupled from the illumination laser. Quantum jumps have also been observed between states of individual molecules[8] and between the cyclotron motional states of a single electron in a Penning trap[4]. As a common feature, all these experiments use fields to probe quantum jumps in matter. Our experiment realizes for the first time the opposite situation, in which the jumps of a field oscillator are revealed via QND measurements performed with matter particles.

We exploit light shifts resolved at the single-photon level, which are experienced by an oscillating dipole in the field of a high-quality-factor (high-Q) cavity. This resolution requires a huge dipole polarizability, which is achieved only with very special systems, such as circular Rydberg atoms[10] or superconducting qubits[11,12] coupled to microwave photons. In our experiment, the measurement of the light shift induced by the field on Rydberg atoms is repeated more than 100 times within the average decay time of individual photons.

The core of the experiment is a photon box (see Fig. 1), which is an open cavity C made up of two superconducting niobium mirrors facing each other (the Fabry–Perot configuration)[13]. The cavity is resonant at 51.1 GHz and cooled to 0.8 K. Its damping time, as measured by the ring-down of a classical injected microwave field, is $T_c = 0.129 \pm 0.003$ s, corresponding to a light travel distance of 39,000 km, folded in the 2.7 cm space between the mirrors. The QND probes are rubidium atoms, prepared in circular Rydberg states[10], travelling along the z direction transverse to the cavity axis. They cross C one at a time, at a rate of 900 s^{-1} with a velocity $v = 250\ m\ s^{-1}$ (see Methods). The cavity C is nearly resonant with the transition between the two circular states e and g (principal quantum numbers 51 and 50, respectively). The position- (z-)dependent atom–field coupling $\Omega(z) = \Omega_0 \exp(-z^2/w^2)$ follows the gaussian profile of the cavity mode (waist $w = 6$ mm). The maximum coupling, $\Omega_0/2\pi = 51$ kHz, is the rate at which the field and the atom located at the cavity centre ($z = 0$) exchange a quantum of energy, when the initially empty cavity is set at resonance with the e–g transition[10].

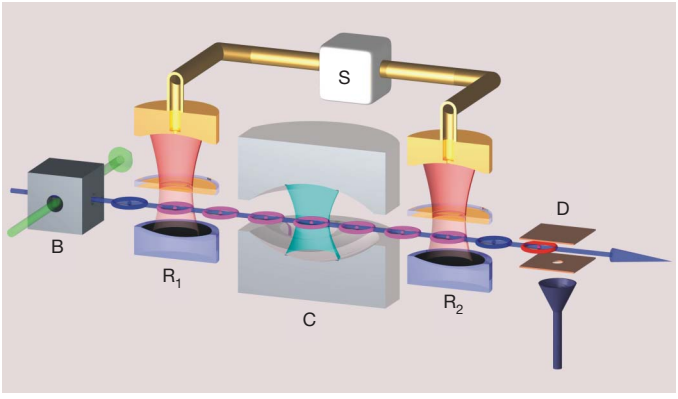


Figure 1 | Experimental set-up. Samples of circular Rydberg atoms are prepared in the circular state g in box B, out of a thermal beam of rubidium atoms, velocity-selected by laser optical pumping. The atoms cross the cavity C sandwiched between the Ramsey cavities R_1 and R_2 fed by the classical microwave source S, before being detected in the state selective field ionization detector D. The R_1–C–R_2 interferometric arrangement, represented here cut by a vertical plane containing the atomic beam, is enclosed in a box at 0.8 K (not shown) that shields it from thermal radiation and static magnetic fields.

[1]Laboratoire Kastler Brossel, Département de Physique de l'Ecole Normale Supérieure, 24 rue Lhomond, 75231 Paris Cedex 05, France. [2]Collège de France, 11 place Marcelin Berthelot, 75231 Paris Cedex 05, France. †Present address: Johannes Gutenberg Universität, Institut für Physik, Staudingerweg 7, 55128 Mainz, Germany.

If the atomic frequency is detuned from the cavity mode by $\delta/2\pi$ with $|\delta| \geq \Omega_0$, emission and absorption of photons by the probe atoms are suppressed owing to the adiabatic variation of $\Omega(z)$ when the atom crosses the gaussian cavity mode (see Methods). The atom–field coupling results in shifts of the atomic and cavity frequencies[9]. The atomic shift depends on the field intensity and thus provides QND information on the photon number n. Following a proposal made in refs 14 and 15, our aim is to read this information by an interferometric method and to monitor the jumps of n between 0 and 1 under the effect of thermal fluctuations and relaxation in the cavity.

Before entering C, the atoms are prepared in a superposition of e and g by a classical resonant field in the auxiliary cavity R_1 (see Fig. 1). During the atom–cavity interaction, this superposition accumulates a phase $\Phi(n,\delta)$. The atomic coherence at the exit of C is probed by subjecting the atoms to a second classical resonant field in R_2, before detecting them in the state-selective counter D. The combination of R_1, R_2 and D is a Ramsey interferometer. The probability of detecting the atom in g is a sine function of the relative phase of the fields in R_1 and R_2. This phase is adjusted so that the atom is ideally found in g if C is empty ($n = 0$). The detuning $\delta/2\pi$ is set at 67 kHz, corresponding to $\Phi(1,\delta) - \Phi(0,\delta) = \pi$. As a result, the atom is found in e if $n = 1$. As long as the probability of finding more than one photon remains negligible, e thus codes for the one-photon state, $|1\rangle$, and g for the vacuum, $|0\rangle$. The probability of finding two photons in a thermal field at $T = 0.8$ K is only 0.3%, and may be neglected in a first approximation.

We first monitor the field fluctuations in C. Figure 2a (top trace) shows a 2.5 s sequence of 2,241 detection events, recording the birth, life and death of a single photon. At first, atoms are predominantly detected in g, showing that C is in $|0\rangle$. A sudden change from g to e in the detection sequence at $t = 1.054$ s reveals a jump of the field intensity, that is, the creation of a thermal photon, which disappears at $t' = 1.530$ s. This photon has survived 0.476 s (3.7 cavity lifetimes), corresponding to a propagation of about 143,000 km between the cavity mirrors.

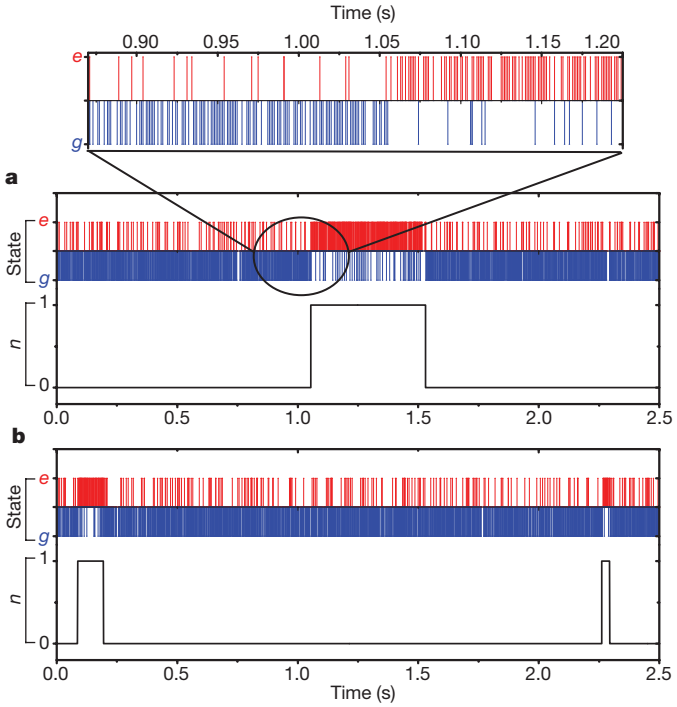


Figure 2 | Birth, life and death of a photon. **a**, QND detection of a single photon. Red and blue bars show the raw signal, a sequence of atoms detected in e or g, respectively (upper trace). The inset zooms into the region where the statistics of the detection events suddenly change, revealing the quantum jump from $|0\rangle$ to $|1\rangle$. The photon number inferred by a majority vote over eight consecutive atoms is shown in the lower trace, revealing the birth, life and death of an exceptionally long lived photon. **b**, Similar signals showing two successive single photons, separated by a long time interval with cavity in vacuum.

The inset in Fig. 2a zooms into the detection sequence between times $t_1 = 0.87$ s and $t_2 = 1.20$ s, and displays more clearly the individual detection events. Imperfections reduce the contrast of the Ramsey fringes to 78%. There is a $p_{g|1} = 13\%$ probability of detecting an atom in g if $n = 1$, and a $p_{e|0} = 9\%$ probability of finding it in e if $n = 0$. Such misleading detection events, not correlated to real photon number jumps, are conspicuous in Fig. 2a and in its inset. To reduce their influence on the inferred n value, we apply a simple error correction scheme. For each atom, n is determined by a majority vote involving this atom and the previous seven atoms (see Methods). The probabilities for erroneous $n = 0$ ($n = 1$) photon number assignments are reduced below 1.4×10^{-3} (2.5×10^{-4}) respectively per detected atom. The average duration of this measurement is 7.8×10^{-3} s, that is, $T_c/17$. The bottom trace in Fig. 2a shows the evolution of the reconstructed photon number. Another field trajectory is presented in Fig. 2b. It displays two single-photon events separated by a 2.069 s time interval during which C remains in vacuum. By probing the field non-destructively in real time, we realize a kind of 'Maxwell demon', sorting out the time intervals during which the thermal fluctuations are vanishing.

Analysing 560 trajectories, we find an average photon number $n_0 = 0.063 \pm 0.005$, slightly larger than $n_t = 0.049 \pm 0.004$, the thermodynamic value at the cavity mirror temperature, 0.80 ± 0.02 K. Attributing the excess photon noise entirely to a residual heating of the field by the atomic beam yields an upper bound to the emission rate per atom of 10^{-4}. This demonstrates the efficient suppression of atomic emission due to the adiabatic variation of the atom–field coupling. This suppression is a key feature that makes possible many repetitions of the QND measurement. Methods based on resonant phase shifts have much larger emission rates, in the 10^{-1} range per atom[3]. Non-resonant methods in which the detector is permanently coupled to the cavity[12] have error rates of the order of Ω_0^2/δ^2, and would require much larger δ/Ω_0 ratios to be compatible with the observation of field quantum jumps.

In a second experiment, we monitor the decay of a single-photon Fock state prepared at the beginning of each sequence. We initialize the field in $|0\rangle$ by first absorbing thermal photons with ~10 atoms prepared in g and tuned to resonance with the cavity mode (residual photon number $\sim 0.003 \pm 0.003$). We then send into the cavity a single atom in e, also resonant with C. Its interaction time is adjusted so that it undergoes half a Rabi oscillation, exits in g and leaves C in $|1\rangle$. The QND probe atoms are then sent across C. Figure 3a shows a typical single photon trajectory (signal inferred by the majority vote) and Fig. 3b–d presents the averages of 5, 15 and 904 such trajectories. The staircase-like feature of single events is progressively smoothed out into an exponential decay, typical of the evolution of a quantum average.

We have neglected so far the probability of finding two photons in C. This is justified, to a good approximation, by the low n_0 value. A precise statistical analysis reveals, however, the small probability of two-photon events, which vanishes only at 0 K. When C is in $|1\rangle$, it decays towards $|0\rangle$ with the rate $(1 + n_0)/T_c$. This rate combines spontaneous $(1/T_c)$ and thermally stimulated (n_0/T_c) photon annihilation. Thermal fluctuations can also drive C into the two-photon state $|2\rangle$ at the rate $2n_0/T_c$ (the factor of 2 is the square of the photon creation operator matrix element between $|1\rangle$ and $|2\rangle$). The total escape rate from $|1\rangle$ is thus $(1 + 3n_0)/T_c$, a fraction $2n_0/(1 + 3n_0) \approx 0.10$ of the quantum jumps out of $|1\rangle$ being actually jumps towards $|2\rangle$.

In this experiment, the detection does not distinguish between $|2\rangle$ and $|0\rangle$. The incremental phase shift $\Phi(2,\delta) - \Phi(1,\delta)$ is 0.88π for $\delta/2\pi = 67$ kHz . The probability of detecting an atom in g when C is in $|2\rangle$ is ideally $[1 - \cos(0.88\pi)]/2 = 0.96$, indistinguishable from 1 within the experimental errors. Since the probability for $n > 2$ is

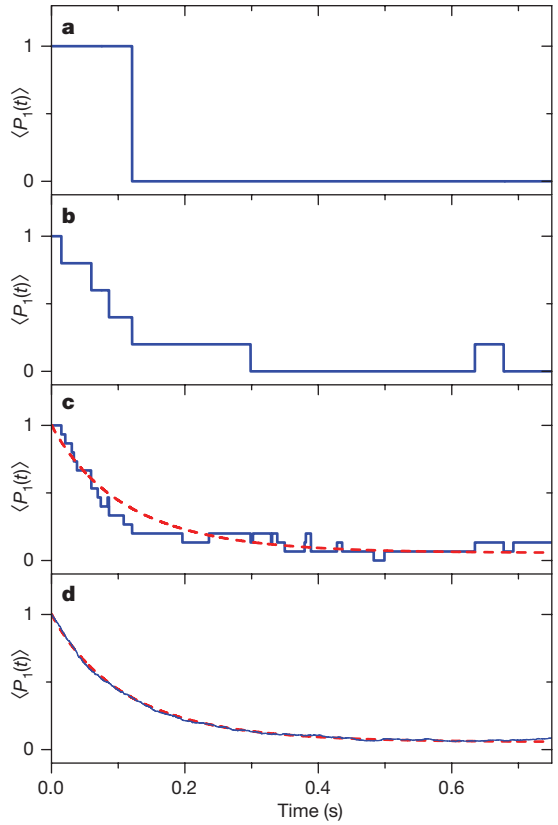


Figure 3 | Decay of the one-photon state. **a**, Measured value of $P_1 = |1\rangle\langle 1|$ as a function of time, in a single experimental realization; **b–d**, averages of 5, 15 and 904 similar quantum trajectories, showing the gradual transition from quantum randomness into a smooth ensemble average. Dotted red line in **c** and **d**, theoretical evolution of the probability of having one photon, $\langle P_1(t)\rangle$, obtained by solving the field master equation with the experimental values of T_c and n_0.

completely negligible, the atoms precisely measure the projector $P_1 = |1\rangle\langle 1|$, *e* (*g*) coding for its eigenvalue 1 (0). Figure 3d thus presents the decay of the ensemble average $\langle P_1(t)\rangle$, that is, the probability of finding one photon in C. The theoretical expectation for $\langle P_1(t)\rangle$ (red dashed line in Fig. 3c and d), obtained by solving the field master equation[9,16] with the known values for T_c and n_0, is nearly indistinguishable from the experimental data in Fig. 3d. Theory predicts—and experiment confirms—for $\langle P_1(t)\rangle$ a quasi-exponential decay with an initial slope corresponding to a time constant $T_c/(1 + 3n_0) = 0.109$ s, slightly shorter than $T_c = 0.129$ s, the damping time of the average photon number.

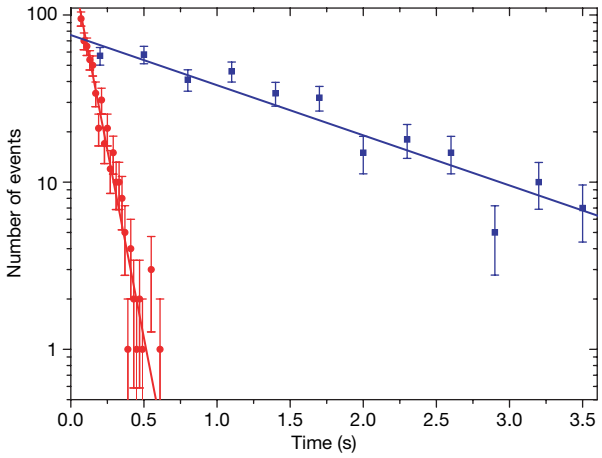


Figure 4 | Lifetime of the one- and zero-photon states. Histograms on a log scale of the durations of the $|1\rangle$ (circles) and $|0\rangle$ (squares) states. The total number of events is 903 for $|1\rangle$ and 338 for $|0\rangle$. The error bars are statistical (± 1 s.d.). The lifetimes $T_1 = 0.097 \pm 0.005$ s and $T_0 = 1.45 \pm 0.12$ s are obtained from the linear fits (solid lines).

Another analysis of the experimental data is provided by Fig. 4, which presents the histograms of the times t of the first quantum jump after preparation of the field at $t = 0$ in $|1\rangle$ (circles) or $|0\rangle$ (squares). The histogram for $|1\rangle$ decays exponentially with the time constant $T_1 = 0.097 \pm 0.005$ s. The small difference from $T_c/(1 + 3n_0)$ is, within error bars, explained by wrong majority votes that can prematurely interrupt a one-photon detection sequence, with a negligible impact on $\langle P_1(t)\rangle$ (see Methods). The vacuum state is prepared by a first QND measurement of the thermal field in C (first vote with majority in *g*). The detection ambiguity between $|0\rangle$ and $|2\rangle$ is then irrelevant. The histogram for $|0\rangle$ also exhibits an exponential decay, with $T_0 = 1.45 \pm 0.12$ s, whereas the expected lifetime is $T_c/n_0 = 2.05 \pm 0.20$ s. The difference is again mainly explained by the rate of false jumps, which most seriously affect the observed lifetime of long-lived states.

The atoms in this QND experiment are witnessing a quantum relaxation process whose dynamics are intrinsically not affected by the measurement. This is fundamentally different from experiments on micromaser bistability, in which the dynamics of a coupled atom–field system exhibit jumps between two stable operating points[17]. Monitoring the photon number quantum jumps realizes an absolute radiation thermometer. The background photon number n_0, extracted from $\langle P_1\rangle$ at thermal equilibrium, explains the field states decay rates well. Even though the two-photon states are not distinguished from the vacuum, their transient appearance with a small probability has an observable effect on the statistics. On single trajectories, however, the ambiguity between $|0\rangle$ and $|2\rangle$ can often be lifted by probabilistic arguments. In Fig. 2, for instance, the long time intervals in which *g* is predominantly detected correspond certainly to vacuum, as their duration is much longer than the lifetime of $|2\rangle$, $T_c/(2 + 5n_0) = 0.057$ s.

This ambiguity is not a fundamental limitation of our QND scheme, which can be extended to monitor larger photon numbers[14,15,18]. By varying the settings of the Ramsey interferometer between probe atoms, we will be able to discriminate between different n values. In the optimal setting[9,18], each detected atom extracts one bit of information about n. Ideally, this quantum analog–digital converter pins down a state with a photon number n between 0 and $N - 1$ using only $\log_2(N)$ atoms. The first QND atom determines in this case the parity of n. Applied to a coherent state, this parity measurement projects the field into a 'Schrödinger cat' state[15,19]. The photon number parity measurement will also allow us to reconstruct the Wigner function of the field in the cavity[20,21] and to follow its time evolution. The decoherence of Schrödinger cat states could be studied in this way[22], providing a direct observation of the evolution from quantum to classical behaviour in a mesoscopic system.

Finally, it is worth noting that, in this QND experiment, a single photon controls the state of a long sequence of atoms. The measurement amounts to a repetitive operation of hundreds of CNOT gates[23] in which the same photon is the control bit (in its $|1\rangle$ or $|0\rangle$ state) and the successive atoms are the targets. This opens promising perspectives for multi-atom entanglement studies.

METHODS

Experimental set-up. The principle of the circular Rydberg atom–microwave cavity set-up is presented in refs 9 and 10. A new superconducting mirror technology has been decisive in reaching very long photon storage times[13]. The mirrors are made of diamond-machined copper, coated with a 12 μm layer of niobium by cathode sputtering. The damping time T_c is two orders of magnitude larger than that of our previous Fabry–Perot cavities made up of massive niobium mirrors[10]. The cavity, whose mirrors have a toroidal surface, sustains two TEM_{900} modes with orthogonal linear polarizations, separated in frequency by 1.2 MHz. The atomic transition is tuned close to resonance with the upper frequency mode by translating a mirror, using piezoelectric actuators. The atoms do not appreciably interact with the other mode. They enter and exit the cavity

through large centimetre-sized ports, avoiding the stray electric fields in the vicinity of metallic surfaces. This ensures a good preservation of the atomic coherence. The Ramsey cavities must have a low Q to minimize enhanced spontaneous emission of the atoms, and yet a well-defined gaussian mode geometry to preclude field leaking into C. To achieve these conflicting requirements, they are made of two parts coupled by a partly reflecting mirror (see Fig. 1). The upper cavity, with $Q = 2 \times 10^3$, defines the mode geometry. It is weakly coupled to the lower one ($Q < 200$), crossed by the atoms.

A QND detection sequence lasting 2.5 s consists of 35,700 atomic sample preparations, separated by 70-μs time intervals. The intensity of the lasers preparing the Rydberg states is kept low enough to limit the occurrence of two or more atoms per sample. This results in most samples being empty. On the average, we detect 0.063 single-atom events per sample. The average atomic detection rate is $r_a = 900\,s^{-1}$.

Each sample undergoes a classical $\pi/2$ pulse of 2-μs duration in R_1 and R_2. The first pulse prepares the atoms in $(|e\rangle + |g\rangle)/\sqrt{2}$. When C contains n photons, the uncoupled atom-cavity states $|e,n\rangle$ and $|g,n\rangle$ evolve into dressed states, shifted respectively, in angular frequency units, by $+([\delta^2 + (n+1)\Omega^2(z)]^{1/2} - \delta)/2$ and $-([\delta^2 + n\Omega^2(z)]^{1/2} - \delta)/2$. The difference between these frequencies, integrated over time ($t = z/v$), yields the phase shift $\Phi(n,\delta)$. Owing to the smooth variation of $\Omega(z)$, the atom–cavity system adiabatically follows the dressed states. The final transition probability between e and g (obtained by numerical integration of the exact Schrödinger equation) is below 10^{-5} for $\delta/2\pi = 67$ kHz. Thus, $(|e\rangle + |g\rangle)/\sqrt{2}$ evolves at the exit of C into $(|e\rangle + \exp[i\Phi(n,\delta)]|g\rangle)/\sqrt{2}$. When $\Phi(1,\delta) - \Phi(0,\delta) = \pi$, the Ramsey pulse in R_2 ideally brings the atom into g if $n = 0$ and into e if $n = 1$.

Majority vote. At each detection time, we determine the photon number by a majority vote, based on the outcomes of the last eight atomic measurements. In case of an equal 4/4 result, we retain the photon number from the preceding vote. This introduces a small hysteresis and reduces the rate of spurious jumps with respect to a simple majority vote with seven or nine atoms. The average duration of this measurement is 7.8×10^{-3} s, resulting in a $\sim 3.9 \times 10^{-3}$ s delay between the occurrence of a quantum jump and its detection. We have determined by numerical simulations that a vote on eight atoms is an optimal trade-off between errors and time resolution. The *a priori* probability of an error in a vote is given by the binomial law. With $p_{g|1} = 13\%$, we erroneously read 0 when there is 1 photon with a probability $\varepsilon_1 \approx (8!/3!5!)(0.13)^5(0.87)^3 = 1.4 \times 10^{-3}$. Similarly $p_{e|0} = 9\%$ results in a false 1 reading with a probability $\varepsilon_0 \approx 2.5 \times 10^{-4}$. These errors are usually corrected after a time of the order of 7.8×10^{-3} s, thus having a negligible impact on ensemble averages such as $\langle P_1(t)\rangle$, which evolve over a much longer timescale. They contribute however to an apparent increase of the $|1\rangle$ and $|0\rangle$ states decay rates. Computing the conditional probability for a vote to be erroneous while all the preceding ones are correct, we find additional decay rates of $0.61\,s^{-1}$ for $|1\rangle$ and $0.12\,s^{-1}$ for $|0\rangle$. Adding these figures to the theoretical decay rates of $|1\rangle$ and $|0\rangle$, we expect to get $T_1 = 0.102 \pm 0.004$ s and $T_0 = 1.64 \pm 0.17$ s.

Received 30 November 2006; accepted 4 January 2007.

1. Braginsky, V. B. & Khalili, F. Y. *Quantum Measurement* (ed. Thorne, K. S.) Chs IV and XI (Cambridge Univ. Press, Cambridge, UK, 1992).
2. Grangier, P., Levenson, J. A. & Poizat, J.-P. Quantum non-demolition measurements in optics. *Nature* **396,** 537–542 (1998).
3. Nogues, G. *et al.* Seeing a single photon without destroying it. *Nature* **400,** 239–242 (1999).
4. Peil, S. & Gabrielse, G. Observing the quantum limit of an electron cyclotron: QND measurements of quantum jumps between Fock states. *Phys. Rev. Lett.* **83,** 1287–1290 (1999).
5. Nagourney, W., Sandberg, J. & Dehmelt, H. Shelved optical electron amplifier: Observation of quantum jumps. *Phys. Rev. Lett.* **56,** 2797–2799 (1986).
6. Sauter, T., Neuhauser, W., Blatt, R. & Toschek, P. E. Observation of quantum jumps. *Phys. Rev. Lett.* **57,** 1696–1698 (1986).
7. Bergquist, J. C., Hulet, R. G., Itano, W. M. & Wineland, D. J. Observation of quantum jumps in a single atom. *Phys. Rev. Lett.* **57,** 1699–1702 (1986).
8. Basché, T., Kummer, S. & Bräuchle, C. Direct spectroscopic observation of quantum jumps of a single molecule. *Nature* **373,** 132–134 (1995).
9. Haroche, S. & Raimond, J. M. *Exploring the Quantum* (Oxford Univ. Press, Oxford, UK, 2006).
10. Raimond, J. M., Brune, M. & Haroche, S. Manipulating quantum entanglement with atoms and photons in a cavity. *Rev. Mod. Phys.* **73,** 565–582 (2001).
11. Schuster, D. I. *et al.* AC-Stark shift and dephasing of a superconducting qubit strongly coupled to a cavity field. *Phys. Rev. Lett.* **94,** 123602 (2005).
12. Schuster, D. I. *et al.* Resolving photon number states in a superconducting circuit. *Nature* **445,** 515–518 (2007).
13. Kuhr, S. *et al.* An ultrahigh Finesse Fabry-Perot superconducting resonator as a photon box for cavity-QED experiments. Preprint at ⟨http://arxiv.org/cond-mat/0612138⟩ (2006).
14. Brune, M., Haroche, S., Lefèvre, V., Raimond, J. M. & Zagury, N. Quantum non-demolition measurement of small photon numbers by Rydberg atom phase-sensitive detection. *Phys. Rev. Lett.* **65,** 976–979 (1990).
15. Brune, M., Haroche, S., Raimond, J. M., Davidovich, L. & Zagury, N. Quantum-non demolition measurements and generation of "Schrödinger cat" states. *Phys. Rev. A* **45,** 5193–5214 (1992).
16. Carmichael, H. J. *An Open System's Approach to Quantum Optics* (Springer, Berlin, 1993).
17. Benson, O., Raithel, G. & Walther, H. Quantum jumps of the micromaser field: Dynamic behavior close to phase transition points. *Phys. Rev. Lett.* **72,** 3506–3509 (1994).
18. Haroche, S., Brune, M. & Raimond, J. M. Measuring photon numbers in a cavity by atomic interferometry: optimizing the convergence procedure. *J. Phys. II France* **2,** 659–670 (1992).
19. Buzek, V. & Knight, P. L. Quantum interference, superposition states of light, and non-classical effects. *Prog. Opt* **34,** 1–158 (1995).
20. Lutterbach, L. G. & Davidovich, L. Method for direct measurement of the Wigner function in cavity QED and ion traps. *Phys. Rev. Lett.* **78,** 2547–2550 (1997).
21. Bertet, P. *et al.* Direct measurement of the Wigner function of a one-photon Fock state in a cavity. *Phys. Rev. Lett.* **89,** 200402 (2002).
22. Davidovich, L., Brune, M., Raimond, J. M. & Haroche, S. Mesoscopic quantum coherences in cavity QED: Preparation and decoherence monitoring schemes. *Phys. Rev. A* **53,** 1295–1309 (1996).
23. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge Univ. Press, Cambridge, UK, 2000).

Acknowledgements We acknowledge funding by ANR, by the Japan Science and Technology Agency (JST), by the EU under the IP projects 'QGATES' and 'SCALA', and by a Marie-Curie fellowship from the European Community (S.K.). We thank P. Bosland, E. Jacques and B. Visentin for the niobium sputtering of the mirrors, and Thomas Keating Ltd for providing radar absorbing material.

Author Contributions S.G. and S.K contributed equally to this work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.B. (brune@lkb.ens.fr).

LETTERS

Adaptive subwavelength control of nano-optical fields

Martin Aeschlimann[1], Michael Bauer[2], Daniela Bayer[1], Tobias Brixner[3], F. Javier García de Abajo[4], Walter Pfeiffer[5], Martin Rohmer[1], Christian Spindler[3] & Felix Steeb[1]

Adaptive shaping of the phase and amplitude of femtosecond laser pulses has been developed into an efficient tool for the directed manipulation of interference phenomena, thus providing coherent control over various quantum-mechanical systems[1–10]. Temporal resolution in the femtosecond or even attosecond range has been demonstrated, but spatial resolution is limited by diffraction to approximately half the wavelength of the light field (that is, several hundred nanometres). Theory has indicated[11,12] that the spatial limitation to coherent control can be overcome with the illumination of nanostructures: the spatial near-field distribution was shown to depend on the linear chirp of an irradiating laser pulse. An extension of this idea to adaptive control, combining multiparameter pulse shaping with a learning algorithm, demonstrated the generation of user-specified optical near-field distributions in an optimal and flexible fashion[13]. Shaping of the polarization of the laser pulse[14,15] provides a particularly efficient and versatile nano-optical manipulation method[16,17]. Here we demonstrate the feasibility of this concept experimentally, by tailoring the optical near field in the vicinity of silver nanostructures through adaptive polarization shaping of femtosecond laser pulses[14,15] and then probing the lateral field distribution by two-photon photoemission electron microscopy[18]. In this combination of adaptive control[1–10] and nano-optics[19], we achieve subwavelength dynamic localization of electromagnetic intensity on the nanometre scale and thus overcome the spatial restrictions of conventional optics. This experimental realization of theoretical suggestions[11–13,16,17,20] opens a number of perspectives in coherent control, nano-optics, nonlinear spectroscopy, and other research fields in which optical investigations are carried out with spatial or temporal resolution.

Figure 1 schematically illustrates our approach to subwavelength optical field control: polarization-shaped ultrashort laser pulses (Methods) illuminate a planar nanostructure, with two-photon photoemission electron microscopy (PEEM)[18] providing the feedback signal from the nanoscale field distribution that is essential for adaptive near-field control. We note that the spatial resolution of two-photon PEEM (~50 nm) is determined by the electron optics and thus independent of the electromagnetic light-field diffraction limit, and that the sensitivity of the two-photon PEEM patterns to the optical field intensities arises from the nonlinear two-photon photoemission process being proportional to the time-integrated fourth power of the local electric-field amplitude (Methods). With these elements in place, a user-specified nanoscopic optical field distribution is realized by processing recorded photoemission patterns in an evolutionary algorithm that directs the iterative optimization of the irradiating laser pulse shape. The basic idea of the experiment is that the measured PEEM pattern identifies the origin of ejected photoelectrons and hence the regions of high local field intensity; controlled variation of the PEEM pattern then proves the spatial control over the nanoscopic field distribution.

The nanostructures we use consist of circular Ag disks with 180 nm diameter and 30 nm height, fabricated by electron-beam lithography on a conductive, 40-nm-thick indium-tin oxide (ITO) film grown on a quartz substrate. The disks are arranged into three dimers that form the arms of a star-like shape (Fig. 1, lower right). The whole nanostructure is about 800 nm across, while the gap between two of the dimer disks is 10 nm wide. After inspection by scanning-electron microscopy (SEM), the sample is mounted in the ultrahigh-vacuum PEEM set-up. The deposition of a small amount of caesium (~0.1 monolayers) reduces the work function of the Ag nanostructure to about 3.1 eV, that is, just below the threshold for two-photon photoemission with 790 nm photons.

Figure 2a shows the photoemission pattern obtained under illumination of the nanostructure with p-polarized femtosecond

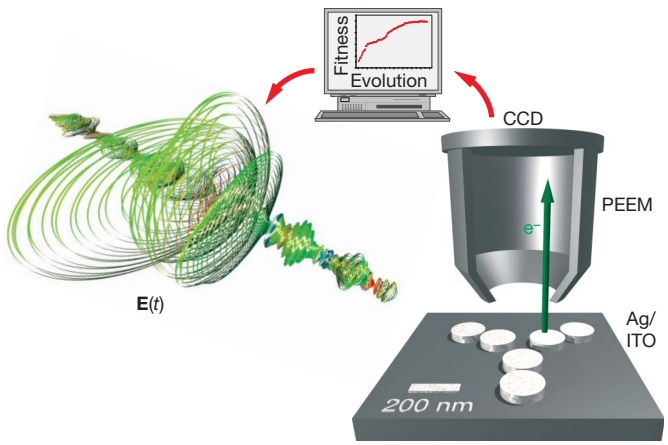


Figure 1 | Experimental scheme. A polarization shaper for ultrashort laser pulses controls the temporal evolution of the vectorial electric field **E**(*t*) on a femtosecond timescale. These pulses illuminate a planar nanostructure in an ultrahigh-vacuum chamber that is equipped with a photoemission electron microscope. The nanostructure consists of six circular Ag islands on an indium-tin oxide (ITO) film and a quartz substrate. A computer-controlled charge-coupled device (CCD) camera records the photoemission image and provides a feedback signal for an evolutionary learning algorithm. Iterative optimization of the pulse-shaper settings leads to an increase in the fitness value and correspondingly allows control over the nano-optical fields.

[1]Fachbereich Physik, Technische Universität Kaiserslautern, Erwin-Schrödinger-Str. 46, 67663 Kaiserslautern, Germany. [2]Institut für Experimentelle und Angewandte Physik, Universität Kiel, Leibnizstr. 19, 24118 Kiel, Germany. [3]Physikalisches Institut, Universität Würzburg, Am Hubland, 97074 Würzburg, Germany. [4]Instituto de Optica-CSIC, Serrano 121, 28006 Madrid, Spain. [5]Fakultät für Physik, Universität Bielefeld, Universitätsstr. 25, 33516 Bielefeld, Germany.

laser pulses (centre wavelength 790 nm, spectral full-width at half-maximum 24 nm, repetition rate 75 MHz, maximum flux 25 μJ cm^{-2}, angle of incidence 65°, projected direction of incidence **k** indicated by the white arrow). The relative geometric position of the nanostructure with respect to the two-photon PEEM image is indicated by white circles; it was established in a separate procedure by comparing the SEM image of a periodic array of nanodisk star structures with the corresponding topographic PEEM image recorded in one-photon photoemission using ultraviolet light (further information concerning this alignment, as well as SEM and ultraviolet-PEEM images, are provided in Supplementary Fig. S1). The two-photon PEEM pattern (Fig. 2a) exhibits three maxima in the photoemission yield, which are located near the inter-disk gaps in the dimers where the near-field enhancement is expected to be strongest[21]. The maximum photoelectron emission is observed from the bottom (downward-oriented) dimer, but significant contributions are seen from all three arms of the nanostructure.

We define two photoelectron emission regions as indicated in Fig. 2a by yellow squares: region A is the sum over two squares covering the upper nanostructure arms, while region B consists of one square covering the lower arm. The integral two-photon photoemission yields from these two areas reflect the local optical field intensities, which allows us to demonstrate adaptive near-field control by first maximizing, and in a second experiment minimizing, the emission ratio A/B through polarization pulse shaping directed by the evolutionary algorithm (Methods). The evolution of the photoemission contrast (A − B)/(A + B) as a function of generation number is shown in Fig. 2b for pulses shaped for maximization (red) and minimization (blue), as well as for unshaped test pulses (black) that monitor experimental fluctuations. The relatively constant level of the test signal indicates stable conditions, while optimization results in a significant increase (or decrease) of the photoemission contrast with respect to unshaped, linearly p-polarized laser pulses.

The PEEM pattern obtained after A/B maximization (Fig. 2c) shows strong emission from the two upper arms of the nanostructure and almost no emission from the bottom part, which had provided the highest count rate when using unshaped laser pulses (Fig. 2a). Analogously, the photoemission after A/B minimization (Fig. 2d) occurs mainly in the lower area while the contribution from the upper two arms is extremely weak. The adaptively determined solution to each optimization problem is robust with respect to slight imperfections in the experimental nanostructures, as discussed in Supplementary Information. These successful optimizations demonstrate that polarization pulse shaping allows adaptive control of the spatial distribution of photoelectrons on a subwavelength scale, and thus of the nanoscopic optical fields that induce photoemission.

The optimally polarization-shaped laser pulses after adaptive A/B maximization and minimization are shown in Fig. 2e and f, respectively, as determined by dual-channel spectral interferometry[14,15]. In this representation, the shape of the quasi-three-dimensional figure indicates the temporal evolution of the polarization state of the electric field, with the colour representing the momentary oscillation frequency[14] (a less intuitive but more quantitative representation is provided and discussed in Supplementary Fig. S2). Contributions from both transverse polarization components are visible in each of the two cases. Whereas the A/B maximization is achieved with a comparatively simple time evolution, the A/B minimization requires a more complex field with varying degrees of ellipticity, orientation and temporal amplitudes.

For comparison purposes, we also recorded PEEM patterns using pulses with simple polarization states characterized by only one pulse-shaping parameter. Here, a uniform spectral phase Φ_0 is assigned for one polarization component while keeping the spectral phase for the second component at zero. If no optical elements were present between the pulse-shaping liquid-crystal display (LCD) and the nanostructure, this would correspond to a simple wave plate with phase retardance Φ_0. In contrast to the adaptive optimization

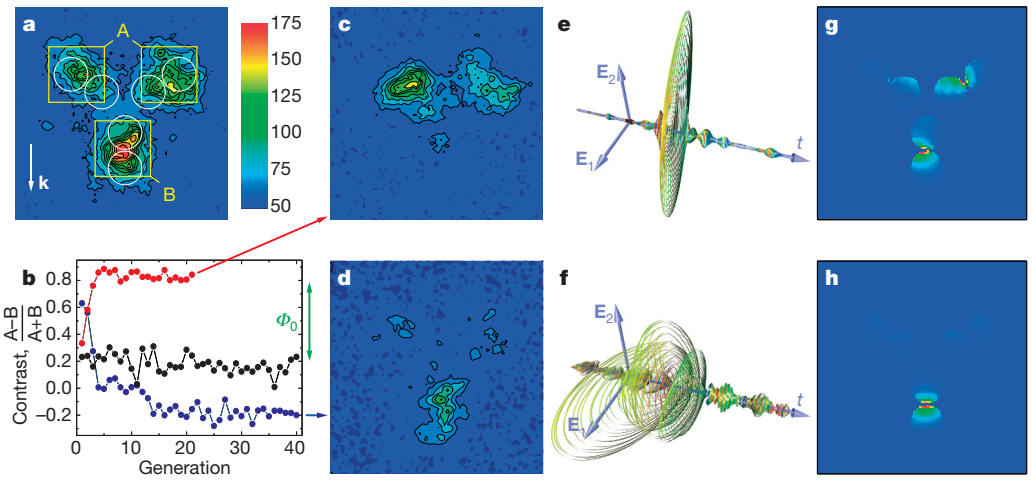


Figure 2 | Nanoscopic field control. a, The experimental photoelectron distribution is displayed on a 1.13 μm × 1.13 μm square for p-polarized excitation of the star-shaped nanostructure with laser pulses at 790 nm. This distribution acts as a reference for the optimization experiments with complex polarization-modulated laser pulses. The colour scale bar represents the photoemission yield in arbitrary units with negligible dark events but including the read-out noise of about 50. The white arrow indicates the projected direction of incidence **k** (incidence angle 65°). The nanostructure positions are marked by white circles, and the regions of interest A and B by yellow squares. **b**, Adaptive optimization of the A/B photoemission ratio leads to increased (red) and decreased (blue) contrast of electron yields from the upper and lower regions as compared to unshaped laser pulses recorded as a reference (black). For comparison, the range of contrast variation achieved by non-adaptive single-parameter scanning of the spectrally constant phase offset between the two pulse-shaper polarization components (green arrow) is smaller and does not provide the same degree of control. **c**, The experimental PEEM image after adaptive A/B maximization using complex polarization-shaped laser pulses (fittest individual of the final generation) shows predominant emission from the upper region. **d**, Photoemission after A/B minimization is concentrated in the lower region. **e**, **f**, The optimal laser pulses as experimentally characterized display complex temporal electric-field evolution for the maximization (**e**) and minimization (**f**) objectives. E_1 and E_2 indicate the two field components that are phase-modulated in the polarization pulse shaper in the first and second LCD layer, respectively. They are at ±45° angles with respect to p-polarization. The overall time window shown is 2 ps. **g**, **h**, The simulated two-photon photoemission pattern for the experimental pulse shapes in the cases of A/B maximization (**g**) and minimization (**h**) qualitatively confirm the experimentally demonstrated nano-optical field control.

experiment, the temporal evolution of polarization and field amplitude is now not adjustable. The contrast variation accompanying the Φ_0 scan is indicated by an arrow in Fig. 2b (see Supplementary Fig. S3 for the complete curve and Supplementary Fig. S4 for a comparison of pulse shapes in Poincaré ellipticity–orientation phase space[22]). Contrast variation is achieved, but the extreme values do not match those reached with adaptive optimization, especially in the case of contrast minimization.

To further substantiate nano-optical field control, we calculated the local electric field for the nanostructure (Fig. 1) and optimized laser pulse shapes (Figs 2e and f) used in the experiments. We solve the frequency-dependent Maxwell equations rigorously by means of the boundary-element method[23,24], taking into account the precise geometry and frequency-dependent dielectric function of the nanostructure, propagation/retardation effects, and the ultrafast time-varying vectorial nature of the incident laser pulse (Methods). This yields the vectorial electric near field as a function of frequency with nanometre spatial resolution, and Fourier transformation delivers the field in the time domain. The two-photon photoemission process is then approximately modelled by calculating the fourth power of the local temporal field one nanometre below the surface of the Ag islands (that is, close to the region from which the photoelectrons actually originate), and integrating over the time period of the laser pulse. Outside the Ag structures, the emission yield is set to zero, as the electron emissivity of the ITO substrate is strongly diminished compared to Ag.

The emission distributions calculated for the optimized laser pulses of Fig. 2e and f are shown in Fig. 2g and h, respectively. Theory predicts photoelectron emission to be concentrated near the gaps between the closely spaced Ag islands, as observed experimentally (Fig. 2a). The larger spatial extent of the experimental distributions as compared to the theoretical ones is due to the limited PEEM resolution of about 50 nm and topographic effects in the photoemission process influencing the emission pattern, whereas the calculation has a grid resolution of 9.3 nm. This is also the reason why only one peak is seen for each of the 'dimer' structures in the experimental PEEM patterns. When comparing predicted and experimentally observed emission patterns, we note that the agreement is very good in the case of A/B minimization (compare Fig. 2d and h) and moderate for the A/B maximization (compare Fig. 2c and g), where the experimentally observed emission occurs almost exclusively from the upper region of the nanostructure whereas the simulation shows a significant contribution also from the bottom part. However, we note that a comparison of the two theoretical plots clearly indicates that the switching from A/B minimization (with bottom-only photoemission, Fig. 2h) to A/B maximization is accompanied by the appearance of photoemission in upper parts of the nanostructure (Fig. 2g); that is, the calculations give the correct tendency in emission changes although the degree of control in the experiment is actually much better. We suspect that the discrepancy between theory and experiment in A/B maximization might arise from our use of a rather simple photoemission model, which uses the fourth power of the local temporal field as a rough estimate for the non-resonant two-photon excitation in an isotropic medium and neglects the influences of varying intermediate-state lifetimes, anisotropic matrix elements, inhomogeneous surface-barrier transmission over the photoelectron energy range, and in particular the role of the surface field in the final photoemission process. For example, the electric-field component normal to the material surface may play a different role from the two tangential components, whereas in the simulation we have considered the complete field. In addition, photoelectrons might also emerge from the 'side walls' of the Ag cylindrical islands, a process not considered in our calculations.

We note that control over photoemission patterns was demonstrated in a recent two-photon PEEM experiment[25] through adjustment of the temporal phase between two unshaped femtosecond excitation pulses, which corresponds to a modification of the excitation spectrum via spectral interference. In contrast, the control mechanism underlying our approach is based on adaptive modification of spectral phase and polarization direction. So far, we have controlled only the spatial degree of freedom of the optical near-field distribution because the feedback signal used for optimization is a time-integrated quantity, but the use of a time-resolved two-photon photoemission signal would readily allow the control of spatial and temporal degrees of freedom simultaneously. Adaptive near-field control is not limited to the present nanostructure. Theoretical investigations of different geometries[13,16,17,20] together with the experimental demonstration here show that this kind of control is broadly applicable.

Summarizing, we have demonstrated mapping of photoemission patterns governed by local optical near fields with nanometre resolution, and shown that optical near fields can be manipulated on a subdiffraction length scale. In effect, we have realized an ultrafast nano-optical switch or multiplexer, delivering 'packets' of electromagnetic energy at desired locations below the diffraction limit. This capability should find many applications, one example being spatial–temporal control of extended quantum systems such as coupled quantum dots or nanowires; in a quantum information processing context, such control could be used to optically and selectively address qubits that are only several nanometres apart. In plasmonic nanodevices, energy transport could be controlled on a nanometre length scale[20]. Near-field imaging techniques[26] and surface-enhanced Raman spectroscopy could also benefit from a subwavelength optimization of electromagnetic fields. The adaptive scheme also provides a femtosecond electron source with a nanoscale-controllable origin. And ultimately, spatial–temporal spectroscopy with direct nanometre and femtosecond resolution[13] might become possible, enabling the direct spatial probing of coupling and energy transfer processes on the nanoscale. Just as the introduction of adaptive femtosecond laser pulse shaping has brought significant advances to the area of coherent chemical reaction control, we expect the concept reported here to provide analogous opportunities in nanophotonics.

METHODS

Polarization pulse shaping. The femtosecond polarization pulse shaper contains a two-layer 128-pixel LCD spatial light modulator in the Fourier plane of a zero-dispersion compressor in folded 4*f* configuration[14,15]. The first and second LCD layers modulate the spectral phase of two transverse polarization components, leading to a spectral variation of polarization state and phase. Hence in the time domain, the intensity, momentary oscillation frequency, and polarization state (that is, elliptical eccentricity and orientation) can all be made to vary within a single laser pulse. In the present improved experimental set-up, we use volume-holographic gratings to achieve high diffraction efficiency for both polarization components, avoiding efficiency-compensating Brewster windows used earlier[14,15]. Experimental characterization of the polarization-shaped laser pulses proceeds directly in the beam path of the experiment via dual-channel spectral interferometry with an unshaped reference laser pulse whose spectral phase is determined separately by SPIDER[27].

Adaptive optimization. In the evolutionary algorithm[28], pulse-shape optimization occurs for both LCD layers independently. For reasons of maintaining experimental stability, the convergence rate is desired to be as high as possible, and thus the complexity of available pulse shapes was reduced by coupling together sequences of four neighbouring LCD pixels. This leads to a reduction of the overall temporal pulse-shaping window to approximately 1 ps, which is still appropriate to control two-photon photoemission without resonant (and longer-lived) intermediate levels. As we are interested in mapping photoemission with nanometre spatial resolution, the issue of long-term drifts of the sample position with respect to the PEEM detection needs to be addressed. After each generation of evolutionary optimization, a PEEM picture with an unshaped laser pulse is recorded as a reference. Performing a two-dimensional cross-correlation with the reference pattern obtained at the beginning of the experiment, the momentary drift vector is calculated automatically and extrapolated linearly to correct for the drift of each individual PEEM picture during the optimization. The good performance of this procedure is apparent from the small variations in the corrected monitoring signal (black curve in Fig. 2b).

Photoemission electron microscopy. The photoemission electron microscope uses an electrostatic lens system to map the lateral variations in the

photoemission yield from the nanostructured surface at ≤50 nm resolution[29]. After amplification using an imaging-quality micro-channel plate, the image is projected onto a screen of phosphorescent material, which is then read out by a Peltier-cooled CCD camera. This signal is used as the input for the evolutionary algorithm for pulse-shape optimization. PEEM in combination with the two-photon photoemission technique is particular sensitive to variations in the local near-field due to the fourth power dependence of the two-photon photoemission yield on the electric field amplitude[30]. The conventional contrast mechanisms of PEEM (work-function variations and topography) do not contribute to the two-photon PEEM signal from the nanostructured surface, as is evident from a direct comparison of the nanoscale structure and the two-photon photoemission yield (Fig. 2a).

Near-field calculation. In the calculations of the near-field response with the boundary-element method[23,24], retardation effects and material properties are taken into account, but the influence of the ITO layer and the quartz substrate is neglected. The vectorial near field is obtained separately for the two cases of s- and p-polarized excitation and for each frequency component of the LCD pulse-shaper pixels. Each response term is then weighted by the (complex-valued) experimentally determined polarization-shaped amplitude-and-phase factor of the incident pulse, and coherent vectorial superposition delivers the full nanoscopic field[13,16]. As only smooth focusing is employed in the experiment ($f = 33$ cm), all simulations are performed in plane-wave approximation, and tight-focusing calculations[17] are not required.

Received 12 September 2006; accepted 3 January 2007.

1. Judson, R. S. & Rabitz, H. Teaching lasers to control molecules. *Phys. Rev. Lett.* **68,** 1500–1503 (1992).
2. Assion, A. *et al.* Control of chemical reactions by feedback-optimized phase-shaped femtosecond laser pulses. *Science* **282,** 919–922 (1998).
3. Meshulach, D. & Silberberg, Y. Coherent quantum control of two-photon transitions by a femtosecond laser pulse. *Nature* **396,** 239–242 (1998).
4. Weinacht, T. C., Ahn, J. & Bucksbaum, P. H. Controlling the shape of a quantum wavefunction. *Nature* **397,** 233–235 (1999).
5. Brixner, T., Damrauer, N. H., Niklaus, P. & Gerber, G. Photoselective adaptive femtosecond quantum control in the liquid phase. *Nature* **414,** 57–60 (2001).
6. Dudovich, N., Oron, D. & Silberberg, Y. Single-pulse coherently controlled nonlinear Raman spectroscopy and microscopy. *Nature* **418,** 512–514 (2002).
7. Herek, J. L., Wohlleben, W., Cogdell, R. J., Zeidler, D. & Motzkus, M. Quantum control of energy flow in light harvesting. *Nature* **417,** 533–535 (2002).
8. Feurer, T., Vaughan, J. C. & Nelson, K. A. Spatiotemporal coherent control of lattice vibrational waves. *Science* **299,** 374–377 (2003).
9. Shapiro, M. & Brumer, P. *Principles of the Quantum Control of Molecular Processes* (Wiley-Interscience, Hoboken, NJ, 2003).
10. Brixner, T. & Gerber, G. Quantum control of gas-phase and liquid-phase femtochemistry. *ChemPhysChem* **4,** 418–438 (2003).
11. Stockman, M. I., Faleev, S. V. & Bergman, D. J. Coherent control of femtosecond energy localization in nanosystems. *Phys. Rev. Lett.* **88,** 067402 (2002).
12. Stockman, M. I., Bergman, D. J. & Kobayashi, T. Coherent control of nanoscale localization of ultrafast optical excitation in nanosystems. *Phys. Rev. B* **69,** 054202 (2004).
13. Brixner, T., García de Abajo, F. J., Schneider, J. & Pfeiffer, W. Nanoscopic ultrafast space-time-resolved spectroscopy. *Phys. Rev. Lett.* **95,** 093901 (2005).
14. Brixner, T. & Gerber, G. Femtosecond polarization pulse shaping. *Opt. Lett.* **26,** 557–559 (2001).
15. Brixner, T., Krampert, G., Niklaus, P. & Gerber, G. Generation and characterization of polarization-shaped femtosecond laser pulses. *Appl. Phys. B* **74,** S133–S144 (2002).
16. Brixner, T., García de Abajo, F. J., Schneider, J., Spindler, C. & Pfeiffer, W. Ultrafast adaptive optical near-field control. *Phys. Rev. B* **73,** 125437 (2006).
17. Brixner, T., García de Abajo, F. J., Spindler, C. & Pfeiffer, W. Adaptive ultrafast nano-optics in a tight focus. *Appl. Phys. B* **84,** 89–95 (2006).
18. Schmidt, O. *et al.* Time-resolved two photon photoemission electron microscopy. *Appl. Phys. B* **74,** 223–227 (2002).
19. Novotny, L. & Hecht, B. *Principles of Nano-Optics* (Cambridge Univ. Press, Cambridge, UK, 2006).
20. Sukharev, M. & Seideman, T. Phase and polarization control as a route to plasmonic nanodevices. *Nano Lett.* **6,** 715–719 (2006).
21. Hao, E. & Schatz, G. C. Electromagnetic fields around silver nanoparticles and dimers. *J. Chem. Phys.* **120,** 357–366 (2004).
22. Brixner, T. Poincaré representation of polarization-shaped femtosecond laser pulses. *Appl. Phys. B* **76,** 531–540 (2003).
23. García de Abajo, F. J. & Howie, A. Relativistic electron energy loss and electron-induced photon emission in inhomogeneous dielectrics. *Phys. Rev. Lett.* **80,** 5180–5183 (1998).
24. García de Abajo, F. J. & Howie, A. Retarded field calculation of electron energy loss in inhomogeneous dielectrics. *Phys. Rev. B* **65,** 115418 (2002).
25. Kubo, A. *et al.* Femtosecond imaging of surface plasmon dynamics in a nanostructured silver film. *Nano Lett.* **5,** 1123–1127 (2005).
26. Sánchez, E. J., Novotny, L. & Xie, X. S. Near-field fluorescence microscopy based on two-photon excitation with metal tips. *Phys. Rev. Lett.* **82,** 4014–4017 (1999).
27. Iaconis, C. & Walmsley, I. A. Spectral phase interferometry for direct electric-field reconstruction of ultrashort optical pulses. *Opt. Lett.* **23,** 792–794 (1998).
28. Baumert, T., Brixner, T., Seyfried, V., Strehle, M. & Gerber, G. Femtosecond pulse shaping by an evolutionary algorithm with feedback. *Appl. Phys. B* **65,** 779–782 (1997).
29. Swiech, W. *et al.* Recent progress in photoemission microscopy with emphasis on chemical and magnetic sensitivity. *J. Electron Spectrosc. Relat. Phenom.* **84,** 171–188 (1997).
30. Cinchetti, M. *et al.* Photoemission electron microscopy as a tool for the investigation of optical near fields. *Phys. Rev. Lett.* **95,** 047601 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank S. Fechner and G. Krampert for constructing the improved pulse shaper, and the Nano-Bio Center at the University of Kaiserslautern for support in the preparation of the nanostructures. This work was supported by the German Science Foundation (DFG) within an Emmy Noether research group (T.B.), by the Graduiertenkolleg 792 (F.S.), and by the Spanish MEC (F.J.G.A.).

Author Contributions The principal investigators of this work are M.B., T.B., F.J.G.A. and W.P.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to T.B. (brixner@physik.uni-wuerzburg.de).

LETTERS

Non-volcanic tremor and low-frequency earthquake swarms

David R. Shelly[1], Gregory C. Beroza[1] & Satoshi Ide[2]

Non-volcanic tremor is a weak, extended duration seismic signal observed episodically on some major faults, often in conjunction with slow slip events[1–4]. Such tremor may hold the key to understanding fundamental processes at the deep roots of faults, and could signal times of accelerated slip and hence increased seismic hazard. The mechanism underlying the generation of tremor and its relationship to aseismic slip are, however, as yet unresolved. Here we demonstrate that tremor beneath Shikoku, Japan, can be explained as a swarm of small, low-frequency earthquakes, each of which occurs as shear faulting on the subduction-zone plate interface. This suggests that tremor and slow slip are different manifestations of a single process.

Tremor is difficult to locate because it lacks the distinct impulsive, body wave arrivals used by traditional earthquake location methods; but occasionally tremor in Japan contains relatively energetic and isolated pulses that have been identified as low-frequency earthquakes (LFEs) by the Japan Meteorological Agency[5]. Compared with nearby ordinary earthquakes, LFEs are enriched in low frequencies (~1–5 Hz) and depleted at higher frequencies. Precise relocations of LFEs beneath western Shikoku reveal that they lie along the dipping subduction interface at depths of 30–35 km (ref. 6). On the basis of their locations and the character of their waveforms, LFEs were inferred to represent fluid-enabled shear slip on the plate boundary as part of concurrently observed slow slip events[6], rather than fluid flow as previously proposed[5].

To test this interpretation further, Ide *et al.* calculated the mechanism of LFEs using two independent methods that exploit waveform similarity between LFEs and regular earthquakes in the subducting slab: analysis of LFE P-wave first motions and an empirical moment tensor inversion of LFE S waves[7]. As shown in Fig. 1, these techniques each yield results consistent with the mechanisms of slow slip events as well as the most recent megathrust earthquake in this area. All these lines of evidence indicate that LFEs are generated by shear slip on the plate interface. Does a similar mechanism generate continuous tremor?

If the same shear-faulting source generates both LFEs and tremor, we might expect to see additional weaker events within tremor with waveforms similar to the previously identified LFEs[6]. The spectrum of tremor tracks that of LFEs, but with slightly smaller amplitude (Fig. 2), which supports the possibility of a common physical mechanism for the two phenomena.

To test this hypothesis, we used the waveforms of 677 of the best-recorded LFEs in this region as 'template events' in a matched-filter technique to search tremor waveforms systematically for portions of the tremor that strongly resemble one or more template LFEs[8]. We require each template event to be recorded at a minimum of six three-component stations. Correlation coefficients from these stations/components are then stacked to produce an 'array' correlation sum. We record a detection when this correlation sum peaks above a threshold value.

As a detected event must possess the same pattern of waveforms across multiple stations and components as the template event, their locations must closely coincide. An example of a positive detection is shown in Fig. 3, which plots the correlation sum as a function of time, as well as waveforms and cross-correlation coefficients at the time of detection across the network. Additional examples of detections and non-detections are shown in Supplementary Fig. S1. Because we are working with complex waveforms with low signal-to-noise ratios, individual channels of data show relatively weak correlations and contain insufficient power to detect events when examined in isolation. The strength of the matched-filter approach comes from simultaneously considering waveforms across the network, which increases the detection power dramatically.

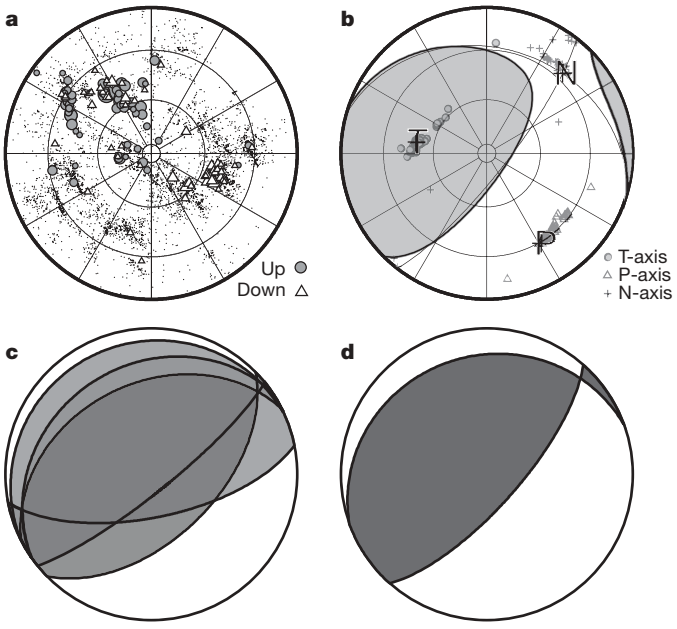


Figure 1 | Comparison of LFE, slow slip event, and megathrust earthquake mechanisms. a, P-wave first motions determined by Ide *et al.* for low frequency earthquakes by cross-correlation-based first motion determination[7]. Filled circles and open triangles indicate compressional (up) and dilatational (down) first motions for LFE P waves, respectively. The signal-to-noise ratio for most observations (small dots) is too low to determine the polarity. **b**, Moment tensor inversion results from empirical Green's function analysis of LFE S waves. Tension (T), pressure (P) and null (N) axes are shown together with symbols showing uncertainty and corresponding P-wave first motion distribution. **c**, Overlay of the mechanism for three slow slip events near the study area[15]. **d**, Mechanism of the 1946 Nankai earthquake[16], which is the most recent megathrust earthquake in this region and representative of relative plate motion between the Philippine Sea plate and the over-riding plate on the dipping plate interface of the Nankai trough subduction zone. All are shown in equal area projection of the lower focal hemisphere.

[1]Department of Geophysics, 397 Panama Mall, Stanford University, Stanford, California 94305-2215, USA. [2]Department of Earth and Planetary Science, University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, Japan.

The ability to extract a weak signal from noisy data is demonstrated by the synthetic example discussed in the Supplementary Methods and shown in Supplementary Fig. S2. We use the correlation sum as our detection statistic, and register a 'strong' detection when a value of 8 times the median absolute deviation is exceeded. For a normally distributed random variable, this corresponds to an exceedence probability of $\sim 3.3 \times 10^{-8}$. On the basis of synthetic tests such as the example shown in Supplementary Fig. S2, we estimate a location uncertainty for 'strong' detections of approximately 3 km.

In order to allow for the possibility that tremor is excited in the vicinity of, but not at exactly the same location as, an LFE in our catalogue, we also allow for 'weak' detections. In this case, we take the maximum correlation coefficient from a 0.4-s window at each station before summing. For the weak detection, the threshold is set at 9 times the median absolute deviation of the distribution of correlation sums plus the median of this distribution, which corresponds to a probability of $\sim 6.4 \times 10^{-10}$, for the gaussian case.

Our detection technique reveals a nearly continuous sequence of LFEs during periods of active tremor. Statistical considerations argue against random detections, but perhaps an even more compelling argument can be made on the basis of the highly clustered nature of the positive detections. The detection statistic that we use is normalized, which means that it does not depend on absolute amplitudes. Therefore, there is nothing in the construction of the measurement to favour positive detections of one LFE over another if the tremor consists of incoherently radiated energy or noise. Thus, we expect that false positive detections should be geographically random.

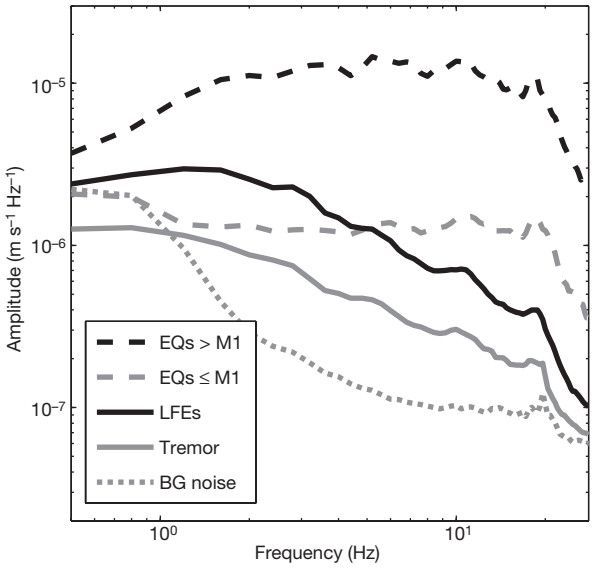


Figure 2 | Comparison of earthquake/LFE/tremor spectra. Shown is a comparison of stacked, uncorrected spectral amplitude from Hi-net velocity seismograms (horizontal components) from multiple nearby events of various kinds. LFEs and regular earthquakes (EQs) are selected from the same hypocentral region (33.4–33.6 °N, 132.6–132.8 °E, depths 25–45 km). Tremor is selected from approximately the same region. For LFEs and regular earthquakes, spectra are calculated for S-wave arrivals for a 2.5-s window, from 0.5 s before to 2.0 s after the catalogue phase arrival. Background (BG) noise is calculated from a 2.5-s window, 2 s before to 0.5 s after the origin time of regular earthquakes. Tremor spectra are calculated from sequential 2.5-s windows over a 400-s time period. Earthquakes are separated by magnitude. We do not divide LFEs by size because their magnitude determination method changed during our study period. In this example, we use: 43 earthquakes with magnitude $M > 1$, 52 earthquakes with $M \leq 1$, and 233 LFEs. LFE (solid black line) and tremor (solid grey line) spectra are highly similar, but tremor has smaller amplitude. Small earthquake (dashed black) and very small (dashed grey) earthquake spectra are highly similar to each other, but lack the rapid decay of amplitude with increasing frequency that is characteristic of both tremor and LFE spectra.

Figure 4 shows that rather than being random, the distribution of positive detections is highly clustered. Not only are the detections spatially coherent during a tremor burst, they also show an interesting time progression. Figure 4 demonstrates the detailed evolution of two 30-min episodes of active tremor on different portions of the plate interface. The second example (Fig. 4b) is notable in that it shows an episode of tremor migration, with the source moving approximately 15 km up-dip along the interface in just over 20 min. This is not a unique occurrence—similar episodes are observed at other times and sometimes propagate in the opposite direction. This along-dip migration rate of $\sim 45\,\mathrm{km\,h^{-1}}$ is much faster than along-strike migration rates of $5–17\,\mathrm{km\,d^{-1}}$ previously reported in this region[1] and in Cacadia[9,10]. Although these slower rates may still govern the longer-term average migration, with the matched-filter technique we can resolve more complex behaviour and faster migration rates along both strike and dip. Notably, we also find instances when multiple tremor sources, separated by up to ~20 km, are active simultaneously. This may explain some of the variation, particularly in depth, found in previous estimates of tremor location in Cascadia[10,11]. A detailed look at tremor behaviour during the periods shown in Fig. 4 is available in the Supplementary Movies.

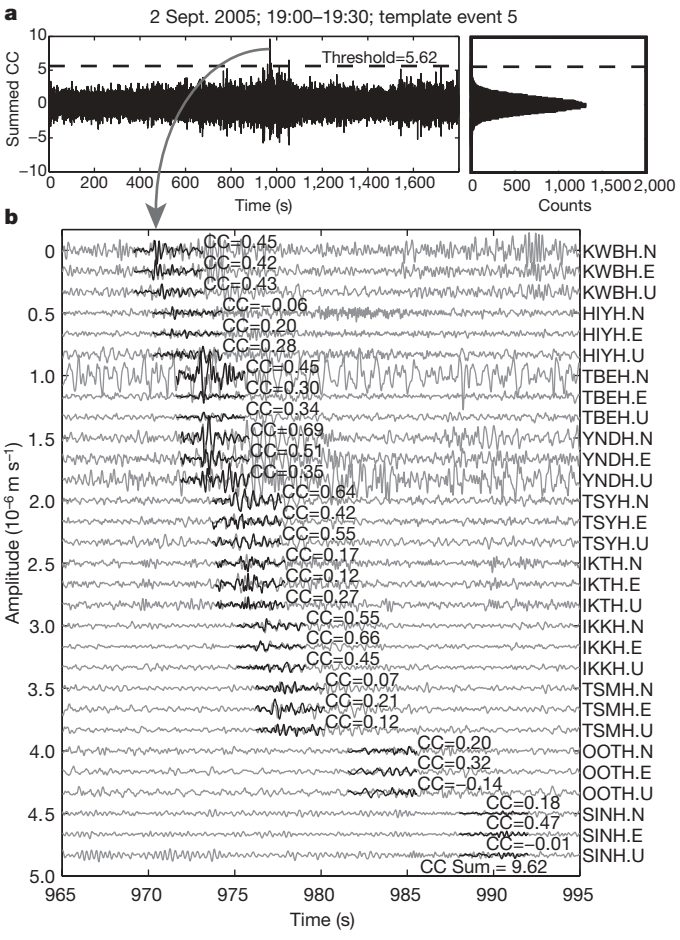


Figure 3 | Example of a detected LFE. a, Correlation sum function for template event 5 during the same 30-min period as highlighted in Fig. 4b. CC, correlation coefficient. A histogram of correlation sum values is shown on the right. **b**, Continuous tremor waveforms are shown in grey and template event waveforms in black for each component of 10 Hi-net stations. The CC for each trace is shown next to the template event waveforms. Station names and components are given to the right of each trace. Additional detections (not shown) are also present during this time window. Waveforms are bandpass filtered between 1 and 8 Hz and template event amplitudes are scaled to match the continuous data. Although individual cross correlation coefficients are modest, they are overwhelmingly positive and extremely unlikely to have occurred by chance.

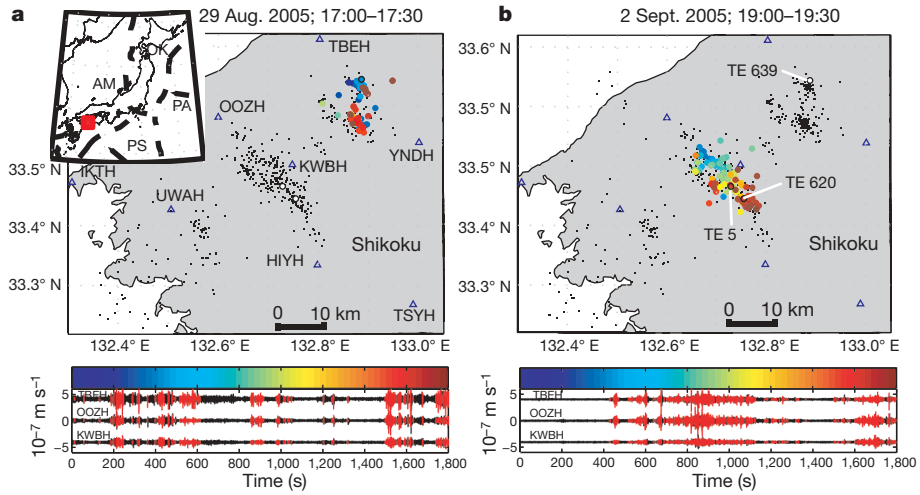


Figure 4 | Detection of LFE swarms forming tremor. a, Top panel; map view of westernmost Shikoku, showing areas active (coloured circles) during the 30-min period on 29 August 2005 beginning at 17:00, colour coded with time (see colour scale below). Only 'strong' detections are shown. Note the clear spatial coherence of detected events with time. The spatial distribution of positive detections is not built into the detection algorithm, but emerges from the data. Black dots show epicentral locations of template LFEs used in this study. The depth of these events corresponds to the plate interface at 30–35 km (ref. 6). Blue triangles denote station locations. Bottom panel shows east-component waveforms at three Hi-net stations, bandpass filtered between 1 and 8 Hz. Portions plotted in red indicate times with a detected event ('strong' or 'weak') similar to a template event. Inset, the regional tectonics, with the red box indicating the area shown in the main panel. PA, Pacific plate; PS, Philippine Sea plate; AM, Amur plate; OK, Okhotsk plate. **b**, Same as **a**, but beginning 2 September 2005 at 19:00. In this episode, a clear up-dip migration of the tremor source can be seen. The locations of template events (TE) referred to in Fig. 3 and Supplementary Fig. S1 are also labelled.

The dense coverage of LFE template sources in the region active in Fig. 4b allows us to attribute the tremor source almost entirely to the sources of known LFEs. This is illustrated by the waveforms shown in Fig. 4b, where portions explained by LFEs are plotted in red. In this case, the match between detected LFEs and tremor is nearly perfect. In the example of Fig. 4a, previously recorded LFEs still explain the vast majority of the tremor, but occasionally weak tremor occurs without LFE detections. This probably reflects the main limitation of our technique—the uneven distribution of our LFE template sources. As we only have template LFEs in places that, at least occasionally, rupture energetically enough to produce an identifiable phase across many stations, some areas of the fault may generate weak tremor without producing 'template-strength' LFEs. In these instances, tremor may be generated too far from a template event to register as a positive detection, even though it may be occurring as a weak LFE. The relatively sparse coverage of the region active in Fig. 4a by template LFE events supports this interpretation.

The heterogeneous distribution of LFEs (Fig. 4) probably reflects properties of the plate boundary. Clusters of relatively strong LFEs may occur in places of geometric or compositional variations where the fault sticks and slips as part of much larger scale slow slip transients—a process analogous to that proposed for some foreshock sequences[12] or earthquake swarms[13,14] in other environments. In this case, high fluid pressure on the plate boundary could allow slip to occur under low shear stress, resulting in relatively slow rupture and slip velocities (compared with ordinary earthquakes) and a corresponding deficit in high frequency energy (Fig. 2).

Using previously recorded and located LFEs as template events, we have established that tremor in Shikoku can be regarded as a swarm of LFEs and thus is generated by a series of small shear slip events on the plate boundary. Our approach allows us to track the source of tremor with unprecedented temporal and spatial resolution, and hence may provide similar precision in monitoring slow slip. It is important to watch such behaviour carefully, as slow slip transients can load adjacent locked portions of a fault and increase the probability of large, damaging earthquakes.

Received 12 October 2006; accepted 6 February 2007.

1. Obara, K. Nonvolcanic deep tremor associated with subduction in southwest Japan. *Science* **296,** 1679–1681 (2002).
2. Rogers, G. & Dragert, H. Episodic tremor and slip on the Cascadia subduction zone: The chatter of silent slip. *Science* **300,** 1942–1943 (2003).
3. Nadeau, R. M. & Dolenc, D. Nonvolcanic tremors deep beneath the San Andreas fault. *Science* **307,** 389 (2005); published online 9 December 2004 (doi:10.1126/science.1107142).
4. Obara, K., Hirose, H., Yamamizu, F. & Kasahara, K. Episodic slow slip events accompanied by non-volcanic tremors in southwest Japan subduction zone. *Geophys. Res. Lett.* **31,** doi:10.1029/2004GL020848 (2004).
5. Katsumata, A. & Kamaya, N. Low-frequency continuous tremor around the Moho discontinuity away from volcanoes in the southwest Japan. *Geophys. Res. Lett.* **30,** doi:10.1029/2002GL015981 (2003).
6. Shelly, D. R., Beroza, G. C., Ide, S. & Nakamula, S. Low-frequency earthquakes in Shikoku, Japan and their relationship to episodic tremor and slip. *Nature* **442,** 188–191 (2006).
7. Ide, S., Shelly, D. R. & Beroza, G. C. Mechanism of deep low frequency earthquakes: further evidence that deep non-volcanic tremor is generated by shear slip on the plate interface. *Geophys. Res. Lett.* **34,** doi:10.1029/2006GL028890 (2007).
8. Gibbons, S. J. & Ringdal, F. The detection of low magnitude seismic events using array-based waveform correlation. *Geophys. J. Int.* **165,** 149–166 (2006).
9. Dragert, H., Wang, K. & Rogers, G. Geodetic and seismic signatures of episodic tremor and slip in the northern Cascadia subduction zone. *Earth Planets Space* **56,** 1143–1150 (2004).
10. Kao, H. *et al.* Spatial-temporal patterns of seismic tremors in northern Cascadia. *J. Geophys. Res.* **111,** doi:10.1029/2005JB003727 (2006).
11. Kao, H. *et al.* A wide depth distribution of seismic tremors along the northern Cascadia margin. *Nature* **436,** 841–844 (2005).
12. Dodge, D. A., Beroza, G. C. & Ellsworth, W. L. Detailed observations of California foreshock sequences: implications for the earthquake initiation process. *J. Geophys. Res.* **101,** 22371–22392 (1996).
13. McGuire, J. J., Boettcher, M. S. & Jordan, T. H. Foreshock sequences and short-term earthquake predictability on East Pacific Rise transform faults. *Nature* **434,** 457–461 (2005).
14. Vidale, J. E. & Shearer, P. M. A survey of 71 earthquake bursts across southern California: Exploring the role of pore fluid pressure fluctuations. *J. Geophys. Res.* **111,** doi:10.1029/2005JB004034 (2006).
15. Hirose, H. & Obara, K. Repeating short- and long-term slow slip events with deep tremor activity, around the Bungo channel region, southwest Japan. *Earth Planets Space* **57,** 961–972 (2005).
16. Ando, M. A fault model of the 1946 Nankaido earthquake derived from tsunami data. *Phys. Earth Planet. Int.* **28,** 320–336 (1982).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This material is based upon work supported by the National Science Foundation. We thank S. Nakamula for assistance with the Hi-net data. All data were obtained from the NIED Hi-net data server.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to D.R.S. (dshelly@pangea.stanford.edu).

Vol 446|15 March 2007|doi:10.1038/nature05615

LETTERS

Evolution and diversity of subduction zones controlled by slab width

W. P. Schellart[1], J. Freeman[1], D. R. Stegman[2], L. Moresi[2] & D. May[2]

Subducting slabs provide the main driving force for plate motion and flow in the Earth's mantle[1–4], and geodynamic, seismic and geochemical studies offer insight into slab dynamics and subduction-induced flow[3–15]. Most previous geodynamic studies treat subduction zones as either infinite in trench-parallel extent[3,5,6] (that is, two-dimensional) or finite in width but fixed in space[7,16]. Subduction zones and their associated slabs are, however, limited in lateral extent (250–7,400 km) and their three-dimensional geometry evolves over time. Here we show that slab width controls two first-order features of plate tectonics—the curvature of subduction zones and their tendency to retreat backwards with time. Using three-dimensional numerical simulations of free subduction, we show that trench migration rate is inversely related to slab width and depends on proximity to a lateral slab edge. These results are consistent with retreat velocities observed globally, with maximum velocities (6–16 cm yr^{-1}) only observed close to slab edges (<1,200 km), whereas far from edges (>2,000 km) retreat velocities are always slow (<2.0 cm yr^{-1}). Models with narrow slabs (≤1,500 km) retreat fast and develop a curved geometry, concave towards the mantle wedge side. Models with slabs intermediate in width (~2,000–3,000 km) are sublinear and retreat more slowly. Models with wide slabs (≥4,000 km) are nearly stationary in the centre and develop a convex geometry, whereas trench retreat increases towards concave-shaped edges. Additionally, we identify periods (5–10 Myr) of slow trench advance at the centre of wide slabs. Such wide-slab behaviour may explain mountain building in the central Andes, as being a consequence of its tectonic setting, far from slab edges.

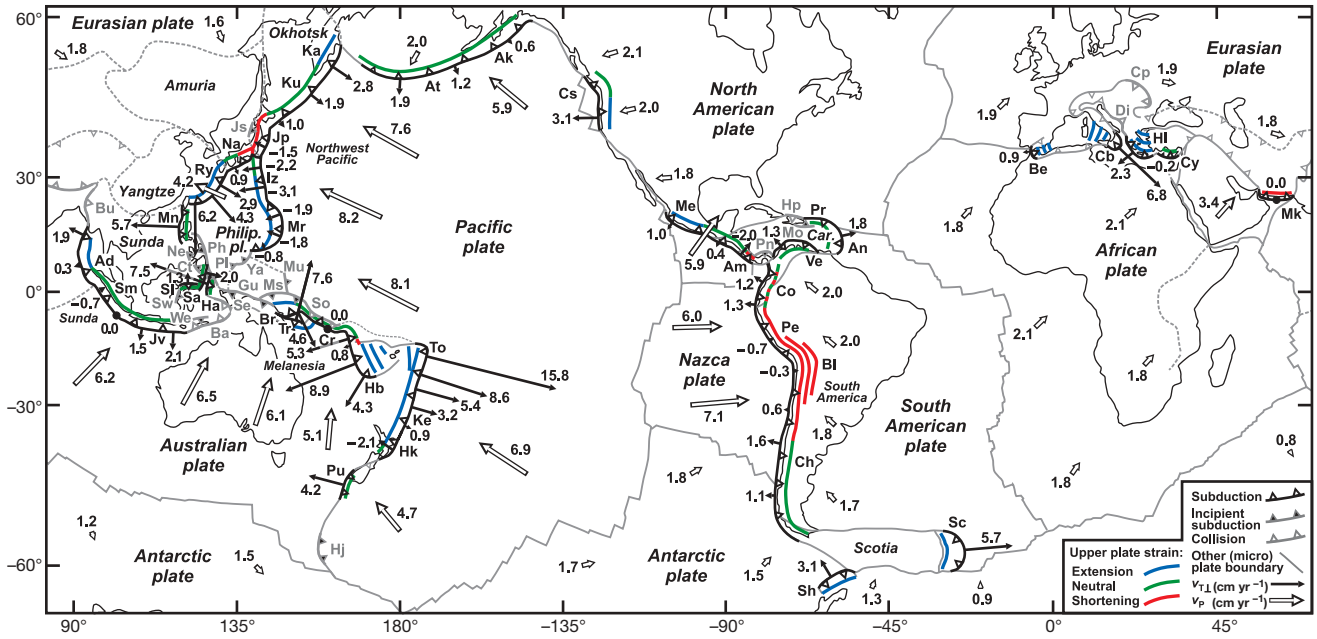


Figure 1 | The major subduction zones on Earth for which the trench-perpendicular trench migration velocity ($v_{T\perp}$) has been calculated. Black arrows illustrate $v_{T\perp}$ vectors, while open arrows illustrate plate velocity (v_P) vectors for 11 major plates. Velocities were calculated using the relative plate motion model from ref. 27 in the Indo-Atlantic hotspot reference frame[28]. Subduction zones: Ad, Andaman; Ak, Alaska; Am, Central America; An, Lesser Antilles; At, Aleutian; Be, Betic Rif; Bl, Bolivia; Br, New Britain; Cb, Calabria; Ch, Chile; Co, Colombia; Cr, San Cristobal; Cs, Cascadia; Cy, Cyprus; Ha, Halmahera; Hb, New Hebrides; Hk, Hikurangi; Hl, Hellenic; Iz, Izu-Bonin; Jp, Japan; Jv, Java; Ka, Kamchatka; Ke, Kermadec; Ku, Kuril; Me, Mexico; Mk, Makran; Mn, Manila; Mr, Mariana; Na, Nankai; Pe, Peru; Pr, Puerto Rico; Pu, Puysegur; Ry, Ryukyu; Sa, Sangihe; Sc, Scotia; Sh, South Shetland; Sl, North Sulawesi; Sm, Sumatra; To, Tonga; Tr, Trobriand; Ve, Venezuela. Collision zones: Ba, Banda; Bu, Burma; Cp, Carpathian; Di, Dinaride; Hp, Hispaniola; Se, Seram; So, Solomon. Incipient subduction zones: Ct, Cotobato; Gu, New Guinea; Hj, Hjort; Js, Japan Sea; Mo, Muertos; Ms, Manus; Mu, Mussau; Ne, Negros; Ph, Philippine; Pl, Palau; Pn, Panama; Sw, West Sulawesi; We, Wetar; Ya, Yap. Incipient subduction zones are young (≤5 Myr), have a short slab (≤150 km) and, together with collision zones, have been excluded from the trench migration calculations in Fig. 2. Trench migration rates for incipient subduction zones are presented in Supplementary Fig. 1.

[1]Research School of Earth Sciences, The Australian National University, Canberra, Australian Capital Territory 0200, Australia. [2]School of Mathematical Sciences, Monash University, Melbourne, Victoria 3800, Australia.

Seismological and geochemical studies suggest that mantle material flows around lateral slab edges towards the mantle wedge[10,12–15]. Such flow is explained as return flow induced by retreat of the trench and rollback of the slab[17,18]. It is expected that lateral slab edges promote rapid rollback by facilitating return flow around the edges instead of underneath the slab tip, as imposed by two-dimensional geometry[17,19]. In nature, the fastest trench-perpendicular trench migration velocities ($v_{T\perp}$ = 6–16 cm yr^{-1}) are indeed observed within 1,200 km of slab edges, whereas at distances >2,000 km $v_{T\perp}$ is <2.0 cm yr^{-1} (Figs 1, 2). Rapid retreat velocities are comparable to fast plate velocities. Figure 2 indicates that proximity to a slab edge is required for, but does not guarantee, rapid trench retreat. Such complexity is expected, because retreat velocity also depends on other factors, including slab density, thickness and length[17], and the rate of tearing between slab and plate at the slab edge, which depends on plate strength[16,20]. These factors generally differ between subduction zones.

Our calculations imply that trench migration is of global importance. For the ~48,800 km of trench considered, 27.6% of the total subduction rate is due to trench migration (mean $v_{T\perp}$ of 1.51 cm yr^{-1} from a total mean convergence of 5.47 cm yr^{-1}). Furthermore, trench migration induces a global upper-mantle volume flux of ~539 km^3 yr^{-1} around the laterally migrating slabs (assuming that $v_{T\perp}$ is constant with depth). This flux expresses itself primarily as toroidal flow[17,21], inducing significant mantle stirring. For comparison, the poloidal mantle flux resulting from trenchward subducting plate motion is only ~193 km^3 yr^{-1} (assuming a constant plate thickness of 100 km).

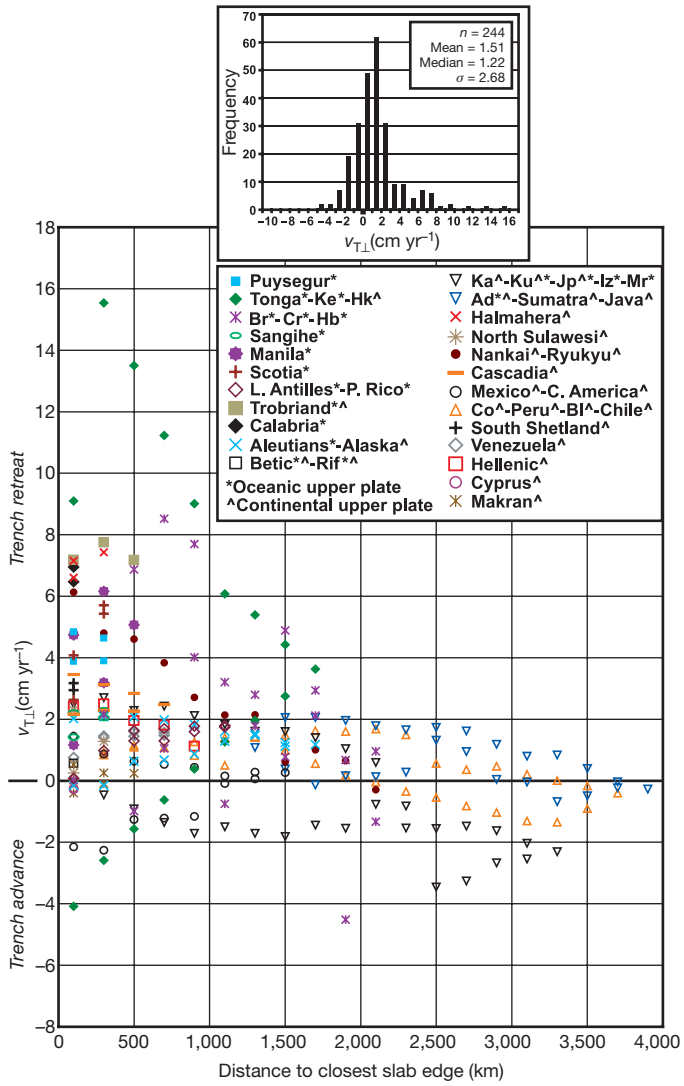


Figure 2 | The proximity of subduction zone segments (200 km wide in trench-parallel extent) to the closest lateral slab edge, plotted against $v_{T\perp}$ for all subduction zones on Earth. Velocities were calculated using the relative plate motion model from ref. 27 in the Indo-Atlantic hotspot reference frame[28]. Positive velocities indicate trench retreat. For details on trench migration calculations see Supplementary Information. Abbreviations as Fig. 1. Inset shows a frequency plot of trench migration velocities for the 244 trench segments investigated.

Recent studies have focused on three-dimensional aspects of subduction zones[17,18,21], but only investigated slab widths ≤1,500 km. We conduct three-dimensional numerical experiments, which for the first time explore the entire range of slab widths (W) found in nature (300–7,000 km). Our experiments are performed with the program 'Underworld' and share the same model design as those of ref. 21, but introduce the important modification of a two-layered plate (a viscoplastic upper half and a viscous lower half). This limits the strength of the upper half, simulating weakening in the brittle upper lithosphere, and allows significant slab–plate coupling through the viscous bottom half to drive plate motion. The subducting plate is laterally homogeneous, thus ignoring trench-parallel buoyancy variation due to aseismic ridges/plateaus and change in lithospheric age, which affect trench geometry and velocity[17,21,22]. Boundary effects have been minimized by implementing widely spaced free-slip boundaries (see Supplementary Information).

Figure 3 shows a wide-slab model (W = 6,000 km) that is representative of simulations with W = 4,000–7,000 km. Initially, trenchward plate motion and trench retreat are slow while the slab dip increases progressively. Trench retreat is accompanied by slab rollback, inducing upper-mantle return flow around the lateral slab edges. Plate and trench velocity increase with time owing to increase in slab length until the slab tip approaches the transition zone. In the centre, trench retreat decelerates and changes to trench advance, while the slab is folded on top of the discontinuity with large recumbent folds (Supplementary Movie 1). Trenchward plate motion remains relatively fast at several cm yr^{-1}. Trench migration in the centre becomes episodic, with phases of retreat and advance ($-1.5 < v_{T\perp} < 1.0$ cm yr^{-1}). Near the slab edges, the trench is predominantly retreating (up to 6.0 cm yr^{-1}, Supplementary Fig. 2), producing a slab-draping geometry on the transition zone.

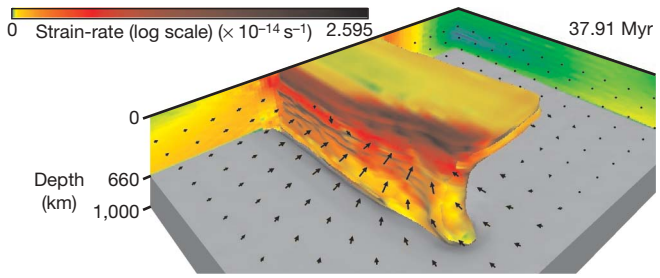


Figure 3 | Image illustrating a wide-slab subduction experiment. (Here W = 6,000 km; see also Supplementary movie 1.) Note that only half of the model is shown (and calculated), because the experiment is symmetrical with respect to a plane through the centre of the subduction zone. Subduction is driven by buoyancy forces only, reflecting natural subduction systems. A 100-km-thick high-viscosity plate (oceanic lithosphere) overlies 900 km of lower-density mantle, subdivided into 560 km of low-viscosity upper mantle and 340 km of high-viscosity lower mantle, confined in a three-dimensional cartesian box. Each experiment started with a 187-km-long slab dipping at 15.5°. Details on the numerical method can be found in Supplementary Information and in ref. 21. The models exclude an overriding plate and elasticity, thereby overestimating slab dip angles[21]. The plate has a viscoplastic upper half and a viscous lower half. The viscosity of the lower half was set to be 200 times the upper-mantle viscosity, producing plate-like behaviour and significant slab–plate coupling to drive plate motion as observed in laboratory simulations[4,17]. Black vectors are located at 200 km depth and illustrate the horizontal flow pattern in the mantle.

Intermediate-width slabs ($W = 2{,}000$–$3{,}000$ km) predominantly retreat at intermediate velocities and produce recumbent folds and minor draping at the transition zone (Supplementary Movie 2). Narrow slabs (300–1,500 km) show the fastest trench retreat velocities but the slowest plate velocities, resulting in slab draping on top of the transition zone (Supplementary Movie 3).

The numerical results show an inverse dependence of $v_{T\perp}$ on W: the wider the slab, the smaller the retreat velocity (Fig. 4a). All models show a local maximum for $v_{T\perp}$ for a slab penetration depth of 400–500 km, that is, before slab tip–transition zone interaction, followed by a decrease in $v_{T\perp}$. For $W \leq 1{,}500$ km, trench retreat dominates ($-0.4 \leq v_{T\perp} \leq 10.0$ cm yr^{-1}), whereas for $W \geq 2{,}000$ km, periods of notable trench advance occur. W also influences $v_{T\perp}$ along the trench itself. For $W \leq 1{,}500$ km, $v_{T\perp}$ is maximum in the centre and decreases towards the edges, thereby forming a curved trench that is concave towards the mantle wedge side (Fig. 4b). For $W \approx 2{,}000$–$3{,}000$ km, $v_{T\perp}$ is of comparable magnitude along the trench, resulting in a more rectilinear trench (Fig. 4c). For $W \geq \sim 4{,}000$ km, $v_{T\perp}$ is minimum in the centre and maximum close to each slab edge (Fig. 4a and d, Supplementary Fig. 2). This results in an overall convex-shaped trench with concave-shaped edges while the slab folds itself around an upper-mantle 'stagnation zone', which experiences minor volume exchange with the mantle wedge. Here, circulation is comparable to two-dimensional models with two poloidal flow cells separated by the slab.

The inverse dependence between W and $v_{T\perp}$ is unexpected, as the face-wise sinking velocity of an oblate ellipsoid through an infinite volume fluid at low Reynolds number ($\ll 1$) depends on the shape of the ellipsoid and the square of the effective diameter[23].

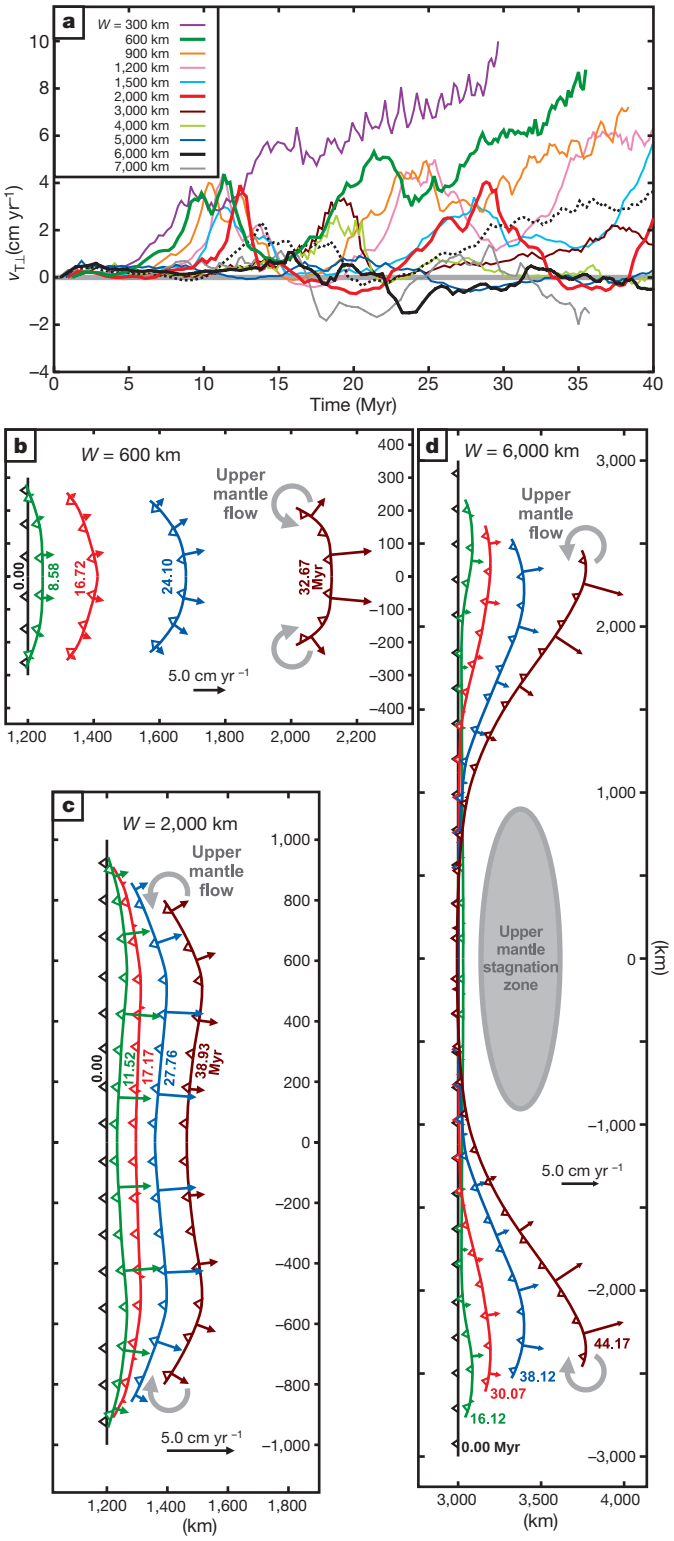


Figure 4 | Influence of W on $v_{T\perp}$ and subduction zone geometry. **a**, Diagram illustrating the influence of W on $v_{T\perp}$ (measured in the centre of the subduction zone; solid lines). For $W = 6{,}000$ km, trench migration velocity (v_T) at 758 km from the slab edge is also plotted (black dotted line). **b**–**d**, Diagrams showing the progressive evolution of the subduction zone for three simulations with a different W. **b**, $W = 600$ km; **c**, $W = 2{,}000$ km; **d**, $W = 6{,}000$ km. Time is given in million years (Myr). Note different length scale in **b**–**d**.

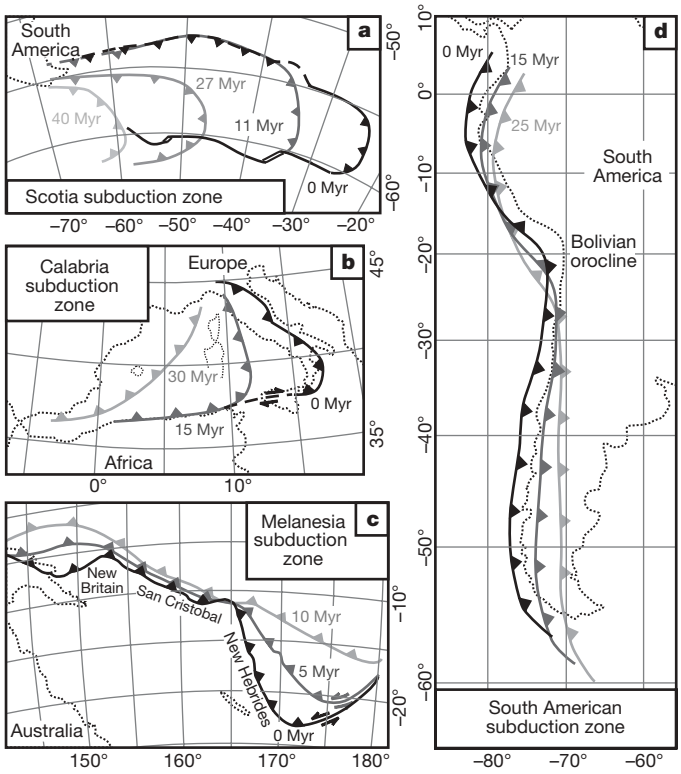


Figure 5 | Progressive evolution of four subduction zones with a different W (measured at the earliest time of the reconstruction). **a**, Evolution of the Scotia subduction zone, $W = 800$ km (modified from ref. 29); **b**, Evolution of the Calabrian subduction zone[11], $W = 1{,}200$ km; **c**, Evolution of the Melanesian subduction zone[20], $W = 4{,}400$ km; **d**, Evolution of the South American subduction zone, $W = 7{,}000$ km, with South American plate motion from ref. 30, adopting a Miocene–Present shortening model for the Andes with a maximum shortening of 350 km in the Bolivian orocline[24] and incorporating accretion/erosion along the trench (see Supplementary Information). Subduction zone migration has been plotted in the following reference frames: **a**, South American plate; **b**, Eurasian plate; **c**, Australian plate; **d**, Indo-Atlantic hotspots. Numerals with Myr indicate time in Myr before the present.

That relationship predicts that the sinking velocity for the widest slab (7,000 km) is 6–25 times faster than for the narrowest slab (300 km), the exact ratio depending on slab length, which suggests that $v_{T\perp}$ increases with increasing slab width. However, the opposite is observed in both nature and experiments. The experiments illustrate that this can be ascribed to the finite upper-mantle thickness that is small compared to W, and complicated rollback-induced flow patterns that are oriented (almost) exclusively around the lateral slab edges (95–100% of the mantle flux)[17,21]. Rollback-induced flow in the upper mantle exerts a viscous drag at the upper–lower mantle boundary that roughly increases with the square of W, while the total slab buoyancy force, which drives sinking and slab retreat, only increases linearly with W. Other complicating factors include slab deformability, allowing variation of $v_{T\perp}$ along the trench.

The models show that rollback-induced return flow produces concave trenches within ~800 km of slab edges, in agreement with the geometry of retreating trench segments near slab edges (Fig. 1). The models further show that subduction zones 300–1,500 km wide become concave, comparable to natural examples with $W < 1{,}600$ km (13 out of 14 are concave), those 2,000–3,000 km wide remain sublinear, comparable to natural examples with $W = 1{,}700$–$3{,}550$ km (4 out of 6 are sublinear), and those 4,000–7,000 km wide become convex, comparable to natural examples with $W = 4{,}400$–$7{,}850$ km (3 out of 4 are convex) (see Supplementary Table 1). The models further agree with the long-term trench kinematics of narrow (Scotia and Calabria, Fig. 5a, b) and wide (Fig. 5c, d) subduction zones, and explain why the four widest subduction zones on Earth (South America, Sunda, Northwest Pacific and Melanesia) are nearly stationary or advancing in the centre.

The Melanesia subduction zone evolved into an overall convex geometry with concave-shaped edges (Fig. 5c). The central San Cristobal segment is nearly stationary with a neutral overriding plate, while the New Britain and New Hebrides segments are retreating, inducing backarc spreading (Manus and North Fiji basins)[20]. The convex Northwest Pacific subduction zone undergoes shortening and trench advance in the centre (Japan), which can be explained by the modelling. Both edges are concave and experience backarc opening (Mariana trough, Central Kamchatka graben) (Fig. 1). Trench advance along the Izu-Bonin-Mariana segment is partly explained by the modelling and partly by rapid westward motion of the overriding Philippine plate and subduction of the Ogasawara plateau and the Caroline ridge. The Sunda subduction zone is not convex, but is nearly stationary in the centre (Sumatra section) and retreating near the northwestern edge (Andaman section with active backarc spreading) and eastern edge (Java section) (Fig. 1). The easternmost Banda section experienced Neogene trench retreat and backarc opening of the Banda Sea before arc–continent collision.

Formation of the Andes mountains and the Bolivian orocline has remained an enigma, as they formed at a subduction zone and not at a continental collision zone such as the Himalayas. A reconstruction of the South American subduction zone shows that it developed a convex geometry with concave edges, whereas the Bolivian orocline has remained nearly stationary for the past 25 Myr (Fig. 5d). The orocline is characterized by compression, orogenesis and shortening in the overriding plate, decreasing progressively northward and southward[24]. Shear wave splitting in the mantle under the South American slab suggests trench-parallel flow in the north and south but negligible trench-parallel flow beneath the orocline[10]. This is consistent with the mantle stagnation zone observed below convex-shaped centres of wide slabs (Fig. 4d). The stagnation zone is capable of supporting large compressive stresses at the subduction interface induced by trenchward overriding plate motion. Indeed, the South American plate is driven westward (Fig. 1) by a combination of forces: suction from the Nazca slab[1,3], slab pull from the Scotia and Lesser Antilles slabs[4,25], and ridge push from the Mid-Atlantic Ridge[25]. Stagnation zone development and intermittent trench advance at the centre of wide slabs can therefore explain enigmatic features of the Central Andes, including mountain building above a subduction zone, the presence of a geoid high below the Bolivian orocline[10] and a rapid increase in palaeo-elevation[26].

Received 8 June 2006; accepted 18 January 2007.

1. Elsasser, W. M. Sea-floor spreading as thermal convection. *J. Geophys. Res.* **76,** 1101–1112 (1971).
2. Forsyth, D. W. & Uyeda, S. On the relative importance of the driving forces of plate motion. *Geophys. J. R. Astron. Soc.* **43,** 163–200 (1975).
3. Zhong, S. & Gurnis, M. Mantle convection with plates and mobile, faulted plate margins. *Science* **267,** 838–843 (1995).
4. Schellart, W. P. Quantifying the net slab pull force as a driving mechanism for plate tectonics. *Geophys. Res. Lett.* **31,** L07611, doi:10.1029/2004GL019528 (2004).
5. Gurnis, M. & Hager, B. H. Controls on the structure of subducted slabs. *Nature* **335,** 317–321 (1988).
6. Christensen, U. R. The influence of trench migration on slab penetration into the lower mantle. *Earth Planet. Sci. Lett.* **140,** 27–39 (1996).
7. Zhong, S. & Gurnis, M. Interaction of weak faults and non-newtonian rheology produces plate tectonics in a 3D model of mantle flow. *Nature* **383,** 245–247 (1996).
8. Kincaid, C. & Olson, P. An experimental study of subduction and slab migration. *J. Geophys. Res.* **92,** 13832–13840 (1987).
9. Griffiths, R. W., Hackney, R. I. & van der Hilst, R. D. A laboratory investigation of effects of trench migration on the descent of subducted slabs. *Earth Planet. Sci. Lett.* **133,** 1–17 (1995).
10. Russo, R. M. & Silver, P. G. Trench-parallel flow beneath the Nazca Plate from seismic anisotropy. *Science* **263,** 1105–1111 (1994).
11. Wortel, M. J. R. & Spakman, W. Subduction and slab detachment in the Mediterranean-Carpathian region. *Science* **290,** 1910–1917 (2000).
12. Peyton, V. *et al.* Mantle flow at a slab edge; seismic anisotropy in the Kamchatka region. *Geophys. Res. Lett.* **28,** 379–382 (2001).
13. Turner, S. & Hawkesworth, C. Using geochemistry to map mantle flow beneath the Lau Basin. *Geology* **26,** 1019–1022 (1998).
14. Gvirtzman, Z. & Nur, A. The formation of Mount Etna as the consequence of slab rollback. *Nature* **401,** 782–785 (1999).
15. Pearce, J. A., Leat, P. T., Barker, P. F. & Millar, I. L. Geochemical tracing of Pacific-to-Atlantic upper mantle flow through the Drake passage. *Nature* **410,** 457–461 (2001).
16. Govers, R. & Wortel, M. J. R. Lithosphere tearing at STEP faults: Response to edges of subduction zones. *Earth Planet. Sci. Lett.* **236,** 505–523 (2005).
17. Schellart, W. P. Kinematics of subduction and subduction-induced flow in the upper mantle. *J. Geophys. Res.* **109,** B07401, doi:10.1029/2004JB002970 (2004).
18. Kincaid, C. & Griffiths, R. W. Laboratory models of the thermal evolution of the mantle during rollback subduction. *Nature* **425,** 58–62 (2003).
19. Dvorkin, J., Nur, A., Mavko, G. & Ben-Avraham, Z. Narrow subducting slabs and the origin of backarc basins. *Tectonophysics* **227,** 63–79 (1993).
20. Schellart, W. P., Lister, G. S. & Toy, V. G. A Late Cretaceous and Cenozoic reconstruction of the Southwest Pacific region: Tectonics controlled by subduction and slab rollback processes. *Earth-Sci. Rev.* **76,** 191–233 (2006).
21. Stegman, D. R., Freeman, J., Schellart, W. P., Moresi, L. & May, D. Influence of trench width on subduction hinge retreat rates in 3-D models of slab rollback. *Geochem. Geophys. Geosyst.* **7,** Q03012, doi:10.1029/2005GC001056 (2006).
22. Vogt, P. R. Subduction and aseismic ridges. *Nature* **241,** 189–191 (1973).
23. Kerr, R. C. & Lister, J. R. The effects of shape on crystal settling and on the rheology of magmas. *J. Geol.* **99,** 457–467 (1991).
24. Kley, J. & Monaldi, C. R. Tectonic shortening and crustal thickness in the Central Andes; how good is the correlation? *Geology* **26,** 723–726 (1998).
25. Coblentz, D. D. & Richardson, R. M. Analysis of the South American intraplate stress field. *J. Geophys. Res.* **101,** 8643–8657 (1996).
26. Ghosh, P., Garzione, C. N. & Eiler, J. M. Rapid uplift of the Altiplano revealed through ^{13}C-^{18}O bonds in paleosol carbonates. *Science* **311,** 511–515 (2006).
27. DeMets, C., Gordon, R. G., Argus, D. F. & Stein, S. Effect of recent revisions to the geomagnetic reversal time scale on estimates of current plate motions. *Geophys. Res. Lett.* **21,** 2191–2194 (1994).
28. O'Neill, C., Müller, D. & Steinberger, B. On the uncertainties in hot spot reconstructions and the significance of moving hot spot reference frames. *Geochem. Geophys. Geosyst.* **6,** Q04003, doi:10.1029/2004GC000784 (2005).
29. Barker, P. F. Scotia Sea regional tectonic evolution: implications for mantle flow and palaeocirculation. *Earth-Sci. Rev.* **55,** 1–39 (2001).
30. Sdrolias, M. & Müller, R. D. Controls on back-arc basin formation. *Geochem. Geophys. Geosyst.* **7,** Q04016, doi:10.1029/2005GC001090 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank O. Oncken, R. Kerr, G. Davies and V. Toy for discussions on subduction processes, mantle dynamics, Andean geology and plate kinematics. We also thank R. Griffiths, M. Sandiford, B. Kennett and D. Müller for providing comments on an early version of the manuscript. Finally, we thank APAC, ACcESS, and VPAC for computational resources and staff from VPAC for technical assistance.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to W.P.S. (wouter.schellart@anu.edu.au).

LETTERS

Hydatellaceae identified as a new branch near the base of the angiosperm phylogenetic tree

Jeffery M. Saarela[1]†, Hardeep S. Rai[1], James A. Doyle[2], Peter K. Endress[3], Sarah Mathews[4], Adam D. Marchant[5], Barbara G. Briggs[5] & Sean W. Graham[1]

Although the relationship of angiosperms to other seed plants remains controversial[1], great progress has been made in identifying the earliest extant splits in flowering-plant phylogeny, with the discovery that the New Caledonian shrub *Amborella trichopoda*, the water lilies (Nymphaeales), and the woody Austrobaileyales constitute a basal grade of lines that diverged before the main radiation in the clade[2–8]. By focusing attention on these ancient lines, this finding has re-written our understanding of angiosperm structural and reproductive biology, physiology, ecology and taxonomy[9–12]. The discovery of a new basal lineage would lead to further re-evaluation of the initial angiosperm radiation, but would also be unexpected, as nearly all of the ~460 flowering-plant families have been surveyed in molecular studies[10]. Here we show that Hydatellaceae, a small family of dwarf aquatics that were formerly interpreted as monocots, are instead a highly modified and previously unrecognized ancient lineage of angiosperms. Molecular phylogenetic analyses of multiple plastid genes and associated noncoding regions from the two genera of Hydatellaceae identify this overlooked family as the sister group of Nymphaeales. This surprising result is further corroborated by evidence from the nuclear gene phytochrome C (*PHYC*), and by numerous morphological characters. This indicates that water lilies are part of a larger lineage that evolved more extreme and diverse modifications for life in an aquatic habitat than previously recognized.

Molecular evidence has been particularly useful in clarifying the phylogenetic positions of groups with highly modified morphologies, such as holoparasites, mycoheterotrophs and certain aquatics. One such group that has resisted placement is Hydatellaceae (Fig. 1), an aquatic family of two genera (*Hydatella* and *Trithuria*), which is restricted to Australasia and India. These genera were traditionally included in Centrolepidaceae, a family of highly reduced monocots. The minute reproductive structures of both families are interpreted as multi-flowered inflorescences surrounded by bracts; individual flowers are unisexual, consisting of a single carpel or stamen, with no associated perianth or bract. Despite these similarities, numerous structural differences warranted separation of Hydatellaceae from Centrolepidaceae[13]. In fact, many features in Hydatellaceae are unknown among graminoid Poales (a clade consisting of Centrolepidaceae, grasses and five other families[14]), including monosulcate pollen, completely anatropous ovules and abundant starchy perisperm (seed storage tissue of nucellar origin). Other characteristics are virtually unknown in monocots; for example, cellular endosperm development, which is restricted to the basal monocot genus *Acorus*. A placement of Hydatellaceae within any of the major clades of monocots has therefore been viewed as problematic[15,16]. Nonetheless, evidence from the plastid gene *rbcL* seemed to confirm conventional views by placing the one species studied (*Trithuria submersa*) in Poales[17,18].

We show here that Hydatellaceae do not belong in Poales, or even in monocots, but instead diverged near the base of angiosperm phylogeny. We re-examined patterns of variation along the published *rbcL* sequence from *T. submersa* as part of a large-scale phylogenetic survey of grasses and relatives, and discovered that it is probably a PCR-based artefact representing a fusion product between a grass and a moss sequence (confirmed by J. Davis, Cornell University, personal communication). Subsequently, we found that combined analysis of multiple plastid genes from *T. submersa* identified it as the sister group of the water lilies (Nymphaeales), with strong bootstrap support from maximum parsimony and maximum likelihood analyses (Fig. 2).

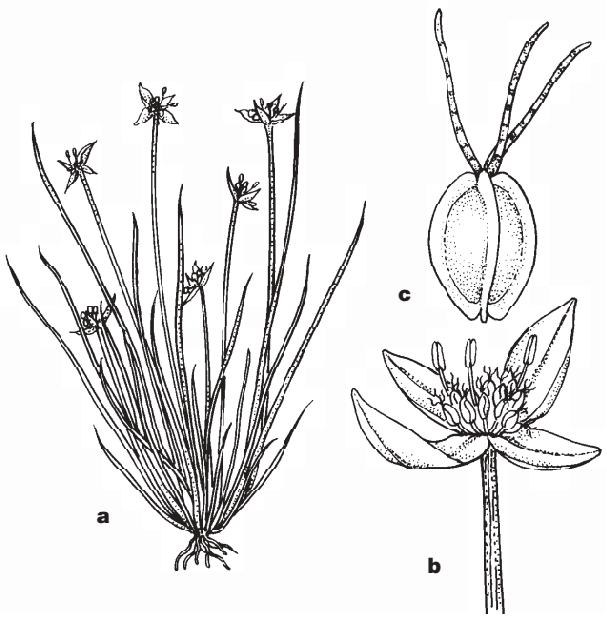


Figure 1 | Hydatellaceae. a, Growth habit of *Trithuria submersa* showing a flowering individual with multiple inflorescence axes (each ~10–35 mm long). **b**, An inflorescence with several staminate flowers and ~10–20 pistillate flowers. The inflorescence is surrounded by an involucre of bracts. **c**, A pistillate flower with several stigmatic hairs, each consisting of a single row of cells[16]. Drawing by N. Oram, reprinted from ref. 31, with permission.

[1]UBC Botanical Garden and Centre for Plant Research (Faculty of Land and Food Systems), Centre for Biodiversity Research, and Department of Botany, 2357 Main Mall, The University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada. [2]Section of Evolution and Ecology, University of California, Davis, California 95616, USA. [3]Institute of Systematic Botany, University of Zurich, Zollikerstrasse 107, 8008 Zurich, Switzerland. [4]The Arnold Arboretum, 22 Divinity Avenue, Harvard University, Cambridge, Massachusetts 02138, USA. [5]Royal Botanic Gardens, Mrs Macquaries Road, Sydney, New South Wales 2000, Australia. †Present address: Research Division, Canadian Museum of Nature, P.O. Box 3443, Station D, Ottawa, Ontario K1P 6P4, Canada.

There is some uncertainty concerning the root of flowering-plant phylogeny[2–8], probably a function of the relatively long branch connecting angiosperms to other seed plants. Because erroneous rooting can lead to mis-inference of ingroup relationships[19], we determined plausible roots for the plastid multigene tree using the Shimodaira–Hasegawa test (see Supplementary Information). Hydatellaceae and Nymphaeales are sister groups in trees rooted at two positions not significantly worse than the optimal one (see arrowheads in Fig. 2), but a root on the *Trithuria* branch is among those rejected. The Hydatellaceae–Nymphaeales relationship is also recovered when substantially more exemplar taxa are included (see Supplementary Information). When each of six plastid data partitions (*atpB*; *ndhF*; *rbcL*; *rpl2*; the 3′-*rps12–trnL*(caa) region; and ten *psb* genes combined) is analysed in separate unrooted analyses (see Methods), we consistently observe a branch separating these two taxa from all other angiosperms, with strong maximum parsimony and maximum likelihood bootstrap support (90–100%).

To corroborate this result using evidence from another genome, we sampled a portion of the nuclear gene phytochrome C (*PHYC*) from *T. submersa*. When analysed with orthologous sequences from other angiosperms that include most of the lineages in Fig. 2, this gene also indicates that Hydatellaceae and water lilies are sister groups, with moderate to strong bootstrap support (70% and 90% in maximum parsimony and maximum likelihood analyses, respectively; Fig. 3a and Supplementary Information). We have also obtained plastid data from *Hydatella*, the other genus in the family, for two noncoding regions that span the plastid transfer RNA genes *trnL*(uaa) and *trnF*(gaa) of *H. inconspicua*. We added these data to an alignment for a somewhat larger plastid region surveyed across the major clades of angiosperms in Fig. 2 and several outgroup taxa[6]. Despite being based on a relatively limited amount of data, phylogenetic analyses again depict Hydatellaceae and water lilies as sister taxa, with moderate bootstrap support (75% and 69% from maximum parsimony and maximum likelihood bootstrap analyses, respectively; Fig. 3b).

Finally, we evaluated whether morphological evidence is consistent with these results by adding published data for Hydatellaceae and two graminoid Poales (Centrolepidaceae and Flagellariaceae) to a morphological matrix for basal angiosperms, including basal eudicots and monocots[9,20] (see Methods). Despite their extensive structural reduction, we could score Hydatellaceae for 64% of the characters in this matrix (see Supplementary Information). Analysis without Hydatellaceae gives some weakly supported results that are strongly over-ruled by molecular data (for example, Nymphaeales are linked with monocots and not located in the basal grade[9]). In most analyses we therefore constrained the arrangement of other taxa to a tree derived from recent molecular and combined morphological–molecular studies (see Methods), allowing the position of Hydatellaceae to be determined by morphology. However, it should be noted that when relationships are not constrained, Hydatellaceae and Nymphaeales form a clade with maximum parsimony bootstrap support comparable to that found with two of our molecular data sets (Fig. 3 and Supplementary Information).

The constrained analysis also indicates that Hydatellaceae and Nymphaeales form a clade (Fig. 4), supported by ten unequivocal synapomorphies (lack of a vascular cambium, lack of pericyclic sclerenchyma, anomocytic stomata, truncate anther connective, boat-shaped pollen, inner integument with two cell layers, palisade exotesta, seed operculum formed by cell enlargement in the inner integument, perisperm and hypogeal germination). Some of these features are among those originally used to segregate Hydatellaceae from Centrolepidaceae[13]. Although most of them (except operculum) occur in other taxa, including various monocots, they are not consistently associated. Of eleven unequivocal synapomorphies of monocots as a

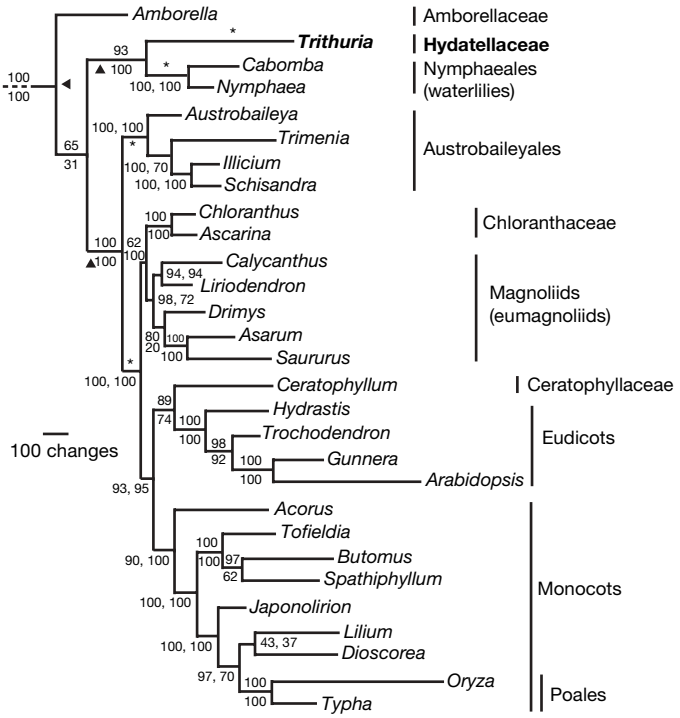


Figure 2 | Phylogenetic placement of *Trithuria submersa* (Hydatellaceae) in angiosperms according to 17 plastid protein-coding loci and six associated noncoding regions. Maximum parsimony and maximum likelihood analyses yield a topology (12,217 steps; $-ln\mathrm{L} = 86{,}798.648$) in which Hydatellaceae are sister to Nymphaeales. Branch lengths are maximum parsimony estimates (ACCTRAN optimization); bootstrap values are noted near branches (maximum parsimony above/left; maximum likelihood below/right); and outgroups (*Cycas* and *Ginkgo*) are trimmed for clarity (and the stem lineage shortened). The optimal and two suboptimal roots (arrowheads) cannot be distinguished from each other in the Shimodaira–Hasegawa test when considering these and four additional roots (asterisks, rejected at $P = 0.001$–0.048). See Supplementary Information for species names.

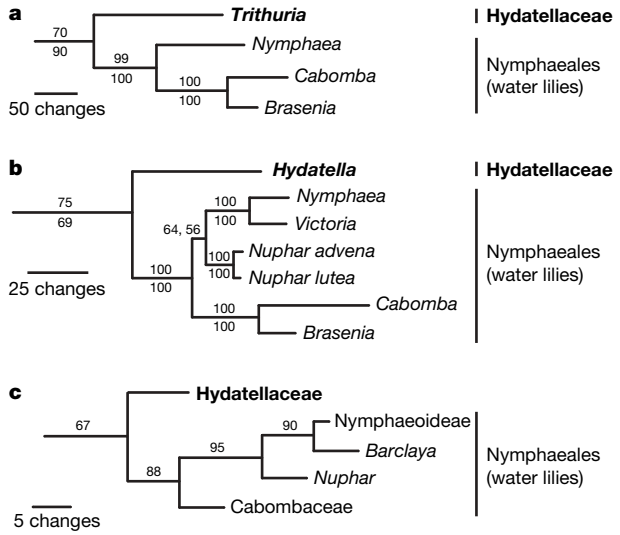


Figure 3 | Local placement of Hydatellaceae according to additional molecular data for *Hydatella* and *Trithuria*, and morphological data for the family as a whole. All analyses support a sister-group relationship with Nymphaeales. **a**, The nuclear gene phytochrome C (*PHYC*) for *Trithuria submersa*. **b**, A portion of the plastid *trnT–F* region for *Hydatella inconspicua*. **c**, An unconstrained morphological analysis. Bootstrap support values are indicated (see Fig. 2); only parsimony bootstrap values are provided for morphology. The *PHYC* and morphological analyses include only angiosperms; the *trnT–F* analysis includes seed-plant outgroups. See Supplementary Information for full trees for angiosperms and species names.

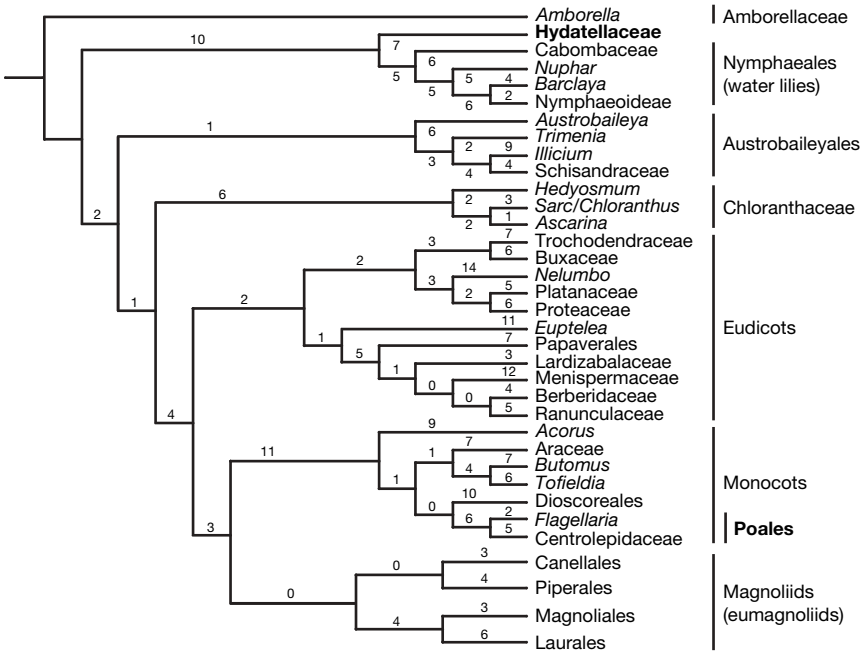


Figure 4 | Most parsimonious position of Hydatellaceae on the basis of morphology. The arrangement of other taxa is constrained to correspond to angiosperm relationships found in recent molecular and combined morphological–molecular studies. Hydatellaceae are the sister group of Nymphaeales. The numbers of unequivocal changes (with respect to the root indicated) are noted near each branch; tree length, 799 steps. For clarity, most phylogenetic structure within magnoliids has been excluded (see ref. 9 for details). *Sarc*, *Sarcandra*.

whole, only three occur in Hydatellaceae (no cambium, boat-shaped pollen and two-layered inner integument), and all of these also occur in Nymphaeales.

In the constrained analyses, a sister-group relationship of Hydatellaceae and Cabombaceae is two steps less parsimonious. The next-best position for Hydatellaceae, four steps less parsimonious than the optimal arrangement, is in monocots, as the sister group of Dioscoreales. Six other positions within monocots, one as the sister group of Centrolepidaceae, are an additional step less parsimonious. Two characters that would favour a relationship of Hydatellaceae with monocots rather than Nymphaeales are P2 sieve tube plastids (which arose independently within Aristolochiaceae) and linear leaves; although the exact point of origin of these characters is equivocal, they each arise twice on the shortest tree. The suboptimal association of Hydatellaceae with Centrolepidaceae would be supported by unisexual flowers, perianth loss and the single carpel, but this relationship would require reversals (or multiple origins) of other derived features of graminoid Poales, including globose, ulcerate pollen, orthotropous ovules and nuclear endosperm development.

Several other characters support a position of Hydatellaceae among the most basal angiosperms. These include completely ascidiate carpels and a four-nucleate embryo sac (the latter needs evaluation in the context of studies among related lineages[11]), features that are considered ancestral in angiosperms[9,11]; if Hydatellaceae and monocots were related, the presence of these characters in Hydatellaceae would represent reversals. Better information on other characters could affect morphological support for inferred relationships. For example, cotyledon number is unknown in Hydatellaceae, which have a 'minute, lens-shaped, incompletely developed embryo'[16], as in some Nymphaeales.

Our results have little effect on previous inferences of the growth habit and ecology of the common ancestor of extant angiosperms[12], because Hydatellaceae are so strongly linked with Nymphaeales, which are also aquatic herbs. A phylogenetic rooting of the flowering plants near *Amborella* and/or water lilies (Nymphaeales) has been robustly supported by numerous analyses of individual and combined sequences from the plastid and other genomes. One contrary result is probably a function of long-branch attraction and low taxon sampling[8]. If *Amborella* alone, or a clade of *Amborella*, Nymphaeales and Hydatellaceae is sister to other angiosperms (see arrowheads in Fig. 2), parsimony optimization implies that the first angiosperms were woody and terrestrial, with Hydatellaceae and Nymphaeales an early line that invaded aquatic habitats. The third plausible rooting, with Hydatellaceae and Nymphaeales sister to all other angiosperms, would imply that the first angiosperms were either woody and terrestrial, or herbaceous and aquatic[12].

It would be misleading to view Hydatellaceae merely as reduced water lilies. First, no member of Nymphaeales approaches the minute, submergence-tolerant, moss-like habit of Hydatellaceae. *Hydatella inconspicua* can grow and flower at >1 m depths; in other species plants are usually initially submerged, but may flower under water or on drying mud at the edges of seasonal pools or swamps. Second, Hydatellaceae have inflorescences rather than solitary flowers, indicating that their common ancestor with Nymphaeales could have had either condition. The possibility that Hydatellaceae are related to the Early Cretaceous aquatic *Archaefructus*[21] should be investigated, as both taxa have inflorescences of naked, unisexual flowers.

Our current knowledge of Hydatellaceae is limited. The family has only recently been discovered in India, with speculation that it has been overlooked elsewhere[16]. Half of the 10 or so species were described in the past 25 yr, and there is substantial morphological variation among them; indeed, more species may await discovery. We lack information on phylogenetic relationships within the family, and it is clear that developmental morphology and monographic work should be high priorities, in addition to studies of the aquatic ecology, biogeography and conservation status of these curious and overlooked basal angiosperms.

METHODS

Molecular analyses. Methods of DNA extraction, amplification, sequencing and alignment follow refs 3 and 22 for plastid data and ref. 4 for *PHYC*. Two sources of DNA were used to generate different *Trithuria submersa* regions: one prepared by the Royal Botanic Gardens, Kew (voucher: J. G. Conran 961 & P. J. Rudall, ADU), the other by K. Bremer, Uppsala University (voucher: Doust *et al.* 1123, MELU). We generated two new *rbcL* sequences using both sources; these are identical for their 483 base pair shared portion (we used the longer of the two (DQ915188, generated from the Conran and Rudall collection) in analyses here;

the shorter *rbcL* sequence (DQ915187) was generated from the Doust *et al.* collection). The problematic published *rbcL* for *T. submersa* (AF458076; refs 14, 17, 18 and 23) was obtained using a different DNA extraction generated from the Doust *et al.* collection; this sequence was not used here. We generated the *trnL–trnF* sequence for *Hydatella inconspicua* from cultivated material sourced from Lake Rotokawau, New Zealand (voucher: P. D. Chapman s.n., NSW accession 428712). All other source details are presented in the Nexus-formatted text files for each data set.

For the plastid-based analyses involving *T. submersa*, we combined protein-coding genes involved in a range of functions, including photosynthesis (*atpB*, *psb* (photosystem II) genes and *rbcL*), chlororespiration (*ndhF*, *ndhB*) and translation (*rpl2*, *rps7*, *rps12*). This multigene plastid matrix includes several conservative noncoding regions (that is, introns in *ndhB*, *rpl2*, and 3′-*rps12*; the intergenic spacers between these genes[3], and between *ndhB* and *trnL*(caa)[22]), but we excluded more rapidly evolving spacer regions in two of the photosystem II gene clusters (*psbB–psbH* and *psbE–psbJ*) from consideration. For analyses of individual regions (subpartitions *atpB*; *ndhF*; *rbcL*; *rpl2*; 3′-*rps12–trnL*(caa) and the combined photosystem II genes), we excluded outgroups (*Cycas* and *Ginkgo*) to minimize the effect of rooting uncertainty on inferred bootstrap support values, as we were interested primarily in the structure of the angiosperm subtree in these analyses. We added the *trnL*(uaa)–*trnF*(gaa) region from *H. inconspicua* to a larger published matrix[6] for *trnT*(ugu)–*trnL*(uaa)–*trnF*(gaa), excluding several mutational hotspots from analysis[6]. We excluded one variable region in the *PHYC* alignment. We performed heuristic maximum-parsimony and maximum-likelihood searches using PAUP* version 4.0b10 (ref. 24) and PHYML version 2.4.4 (ref. 25), respectively, using settings described in the Supplementary Information.

Morphological analyses. For morphological analyses, we added consensuses of *Hydatella* and *Trithuria*, Centrolepidaceae, and Flagellariaceae to a published data matrix[9,20], with modifications described in Supplementary Information; scoring of most characters was based on refs 13, 15, 16, 18 and 26–28, with additional data from primary sources provided in Supplementary Information. All heuristic maximum parsimony searches were performed using PAUP*. We used two backbone constraint trees to fix relationships of all taxa but Hydatellaceae: one corresponding to the maximum parsimony tree from the combined morphological and molecular analyses of refs 9 and 20, the second (see Fig. 4) with modifications on the basis of increasingly robust molecular evidence on relationships in (eu)magnoliids[5,7] (Piperales, Canellales, Magnoliales, Laurales), eudicots[29] and monocots[14,22,23] (see also Fig. 2). The results reported above are based on the second backbone, but we found similar results using the first topology[9,20]. We also evaluated less than optimal arrangements of Hydatellaceae by searching for trees up to six steps less parsimonious than the optimal tree length, and by moving taxa manually with MacClade v. 4.06 (ref. 30). Heuristic maximum parsimony searches (including bootstrap analysis of the unconstrained data) otherwise followed settings described for the molecular analyses. We used the 'Trace all changes' tool in MacClade to determine unequivocal synapomorphies.

Received 13 September 2006; accepted 18 January 2007.

1. Burleigh, J. G. & Mathews, S. Phylogenetic signal in nucleotide data from seed plants: implications for resolving the seed plant tree of life. *Am. J. Bot.* **91,** 1599–1613 (2004).
2. Soltis, P. S., Soltis, D. E. & Chase, M. W. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature* **402,** 402–404 (1999).
3. Graham, S. W. & Olmstead, R. G. Utility of 17 chloroplast genes for inferring the phylogeny of the basal angiosperms. *Am. J. Bot.* **87,** 1712–1730 (2000).
4. Mathews, S. & Donoghue, M. J. Basal angiosperm phylogeny inferred from duplicate phytochromes A and C. *Int. J. Plant Sci.* **161** (6 Suppl.), S41–S55 (2000).
5. Zanis, M. J., Soltis, D. E., Soltis, P. S., Mathews, S. & Donoghue, M. J. The root of angiosperms revisited. *Proc. Natl Acad. Sci. USA* **99,** 6848–6853 (2002).
6. Borsch, T. *et al.* Noncoding plastid *trnT–trnF* sequences reveal a well resolved phylogeny of basal angiosperms. *J. Evol. Biol.* **16,** 558–576 (2003).
7. Qiu, Y.-L. *et al.* Phylogenetic analyses of basal angiosperms based on nine plastid, mitochondrial, and nuclear genes. *Int. J. Plant Sci.* **166,** 815–842 (2005).
8. Leebens-Mack, J. *et al.* Identifying the basal angiosperm node in chloroplast genome phylogenies: sampling one's way out of the Felsenstein zone. *Mol. Biol. Evol.* **22,** 1948–1963 (2005).
9. Doyle, J. A. & Endress, P. K. Morphological phylogenetic analysis of basal angiosperms: comparison and combination with molecular data. *Int. J. Plant Sci.* **161** (6 Suppl.), S121–S153 (2000).
10. Angiosperm Phylogeny Group (APG II). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants. *Bot. J. Linn. Soc.* **141,** 399–436 (2003).
11. Williams, J. H. & Friedman, W. E. The four-celled female gametophyte of *Illicium* (Illiciaceae; Austrobaileyales): implications for understanding the origin and early evolution of monocots, eumagnoliids, and eudicots. *Am. J. Bot.* **91,** 332–351 (2004).
12. Feild, T. S., Arens, N. C., Doyle, J. A., Dawson, T. E. & Donoghue, M. J. Dark and disturbed: a new image of early angiosperm ecology. *Paleobiology* **30,** 82–107 (2004).
13. Hamann, U. Hydatellaceae—a new family of Monocotyledoneae. *N. Zeal. J. Bot.* **14,** 193–196 (1976).
14. Bremer, K. Gondwanan evolution of the grass alliance of families (Poales). *Evolution* **56,** 1374–1387 (2002).
15. Dahlgren, R. M. T., Clifford, H. T. & Yeo, P. F. *The Families of the Monocotyledons: Structure, Evolution, and Taxonomy* (Springer, Berlin, 1985).
16. Hamann, U. in *The Families and Genera of Vascular Plants IV. Flowering Plants. Monocotyledons. Alismatanae and Commelinanae (except Gramineae)* (ed. Kubitzki, K.) 231–234 (Springer, Berlin, 1998).
17. Stevenson, D. W. *et al.* in *Monocots: Systematics and Evolution* (eds Wilson, K. L. & Morrison, D. A.) 17–24 (CSIRO, Collingwood, Australia, 2000).
18. Michelangeli, F. A., Davis, J. I. & Stevenson, D. W. Phylogenetic relationships among Poaceae and related families as inferred from morphology, inversions in the plastid genome, and sequence data from the mitochondrial and plastid genomes. *Am. J. Bot.* **90,** 93–106 (2003).
19. Graham, S. W., Olmstead, R. G. & Barrett, S. C. H. Rooting phylogenetic trees with distant outgroups: a case study from the commelinoid monocots. *Mol. Biol. Evol.* **19,** 1769–1781 (2002).
20. Doyle, J. A. Early evolution of angiosperm pollen as inferred from molecular and morphological phylogenetic analyses. *Grana* **44,** 227–251 (2005).
21. Sun, G. *et al.* Archaefructaceae, a new basal angiosperm family. *Science* **296,** 899–904 (2002).
22. Graham, S. W. *et al.* in *Monocots: Comparative Biology and Evolution (excluding Poales)* (eds Columbus, J. T., Friar, E. A., Porter, J. M., Prince, L. M. & Simpson, M. G.) 3–21 (Rancho Santa Ana Botanic Garden, Claremont, California, 2006).
23. Chase, M. W. *et al.* in *Monocots: Comparative Biology and Evolution (excluding Poales)* (eds Columbus, J. T., Friar, E. A., Porter, J. M., Prince, L. M. & Simpson, M. G.) 63–75 (Rancho Santa Ana Botanic Garden, Claremont, California, 2006).
24. Swofford, D. L. *Phylogenetic Analysis Using Parsimony* (PAUP*)* (Sinauer Associates, Sunderland, Massachusetts, 2002).
25. Guindon, S. & Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52,** 696–704 (2003).
26. Hamann, U. Neue Untersuchungen zur Embryologie und Systematik der Centrolepidaceae. *Bot. Jahrb. Syst.* **96,** 154–191 (1975).
27. Cooke, D. A. in *The Families and Genera of Vascular Plants IV. Flowering Plants. Monocotyledons. Alismatanae and Commelinanae (except Gramineae)* (ed. Kubitzki, K.) 106–109 (Springer-Verlag, Berlin, 1998).
28. Appel, O. & Bayer, C. in *The Families and Genera of Vascular Plants IV. Flowering Plants. Monocotyledons. Alismatanae and Commelinanae (except Gramineae)* (ed. Kubitzki, K.) 208–211 (Springer, Berlin, 1998).
29. Kim, S., Soltis, D. E., Soltis, P. S., Zanis, M. J. & Suh, Y. Phylogenetic relationships among early-diverging eudicots based on four genes: were the eudicots ancestrally woody? *Mol. Phylog. Evol.* **31,** 16–30 (2004).
30. Maddison, D. R. & Maddison, W. P. *MacClade 4: Analysis of Phylogeny and Character Evolution, Version 4.03* (Sinauer Associates, Sunderland, Massachusetts, 2001).
31. Harden, G. J. (ed.) *Flora of New South Wales*. Vol. 4 (Univ. of New South Wales, Kensington, New South Wales, Australia, 1993).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to K. Bremer (Uppsala University) and the Royal Botanic Gardens, Kew, for generously providing DNAs, and to J. Conran, J. Davis, A. Doust, P. Rudall and D. Stevenson and other workers responsible for making the field collections and generating cultivated material. We acknowledge critical review of the manuscript by S. C. H. Barrett, M. W. Chase, T. S. Feild and E. M. Friis. This research was supported by an NSERC Discovery Grant to S.W.G., NSERC postgraduate scholarships to J.M.S. and H.S.R., Alberta Ingenuity and University Graduate Fellowship (University of British Columbia) funding to J.M.S, an NSF grant to S.M, and Royal Botanic Gardens Trust Sydney funding to A.D.M. and B.G.B.

Author Contributions Plastid data were generated by J.M.S., H.S.R. and A.D.M.; nuclear data were generated by S.M.; morphological data were compiled and scored by J.A.D, P.K.E. and B.G.B. Analyses were conceived and performed by S.W.G, J.A.D., J.M.S., H.S.R and S.M. All authors contributed to the writing, which was coordinated by S.W.G. and J.A.D.

Author Information Novel sequences for this study have GenBank accession numbers as follows: *Aphelia brizula* (EF153935, EF153937, EF153939, EF153942, EF153945, EF153948, EF153950, EF153952, EF153954); *Brasenia schreberi* (DQ981792); *Centrolepis monogyna* (EF153934, EF153936, EF153938, EF153941, EF153944, EF153947, EF153949, EF153951, EF153953); *Hydatella inconspicua* (DQ916291); *Trithuria submersa* (AJ419142, DQ915185-DQ915189, DQ981794, EF153940, EF153943, EF153946); *Schisandra sphenanthera* (DQ981793). Alignments used are available for download in Supplementary Information. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to S.W.G. (swgraham@interchange.ubc.ca).

LETTERS

A recurrent mutation in *PALB2* in Finnish cancer families

Hannele Erkko[1]*, Bing Xia[5]*, Jenni Nikkilä[1], Johanna Schleutker[6], Kirsi Syrjäkoski[6], Arto Mannermaa[7], Anne Kallioniemi[6], Katri Pylkäs[1], Sanna-Maria Karppinen[1], Katrin Rapakko[1], Alexander Miron[5], Qing Sheng[5], Guilan Li[5], Henna Mattila[6], Daphne W. Bell[8]†, Daniel A. Haber[8], Mervi Grip[2], Mervi Reiman[1], Arja Jukkola-Vuorinen[3], Aki Mustonen[1], Juha Kere[9,10], Lauri A. Aaltonen[9], Veli-Matti Kosma[7], Vesa Kataja[11], Ylermi Soini[4], Ronny I. Drapkin[5], David M. Livingston[5] & Robert Winqvist[1]

***BRCA1*, *BRCA2* and other known susceptibility genes account for less than half of the detectable hereditary predisposition to breast cancer[1–3]. Other relevant genes therefore remain to be discovered. Recently a new BRCA2-binding protein, PALB2, was identified[4]. The BRCA2–PALB2 interaction is crucial for certain key BRCA2 DNA damage response functions as well as its tumour suppression activity[4]. Here we show, by screening for *PALB2* mutations in Finland that a frameshift mutation, c.1592delT, is present at significantly elevated frequency in familial breast cancer cases compared with ancestry-matched population controls. The truncated PALB2 protein caused by this mutation retained little BRCA2-binding capacity and was deficient in homologous recombination and crosslink repair. Further screening of c.1592delT in unselected breast cancer individuals revealed a roughly fourfold enrichment of this mutation in patients compared with controls. Most of the mutation-positive unselected cases had a familial pattern of disease development. In addition, one multigenerational prostate cancer family that segregated the c.1592delT truncation allele was observed. These results indicate that *PALB2* is a breast cancer susceptibility gene that, in a suitably mutant form, may also contribute to familial prostate cancer development.**

BRCA2, one of the two major breast cancer susceptibility proteins, primarily functions in homologous recombination (HR) and HR-based DNA double-strand break repair (DSBR)[5]. The lifetime breast cancer risk of heterozygous *BRCA2* mutation carriers is about 60–85%[1,6]. In addition, BRCA2 is also a Fanconi anaemia protein[7]. Recently, a previously unidentified BRCA2 binding factor, PALB2 (for 'partner and localizer of BRCA2'), was identified and shown to be crucial for the association of BRCA2 with chromatin and nuclear structures and for its DNA damage response functions[4]. Importantly, certain breast-cancer-associated missense variants in the PALB2-binding domain of BRCA2 abrogate BRCA2–PALB2 complex formation, and these BRCA2 variants are defective in HR-based DSBR, emphasizing the importance of proper interplay between PALB2 and BRCA2 for this essential function[4]. Given the above-noted discoveries and the dependence of BRCA2 function on PALB2, we examined whether *PALB2* is a gene predisposing susceptibility to hereditary breast cancer in its own right.

To explore this possibility, we screened for germline mutations in the exonic regions and splice junctions of the *PALB2* gene, in 113 *BRCA1*/*BRCA2* mutation-negative breast or breast–ovarian cancer families from northern Finland. As shown in Table 1, a total of six different exonic variant alleles were identified in affected index individuals. Four of these changes were also detected at similar frequencies in the control population, suggesting that they are not cancer-associated. This view was supported by the results obtained from computer simulations using PolyPhen, ESEfinder and NNSplice software. By contrast, one alteration (c.1592delT) was detected in three (2.7%) index individuals, but only in six (0.2%) of 2,501 controls ($P = 0.005$; odds ratio (OR) 11.3; 95% confidence interval (CI) 1.8–57.8), therefore suggesting a significant disease association. This alteration should result in a frame-shift at Leu 531, with the new reading frame progressing for 28 codons before termination. Another alteration, 3433G→C (G1145R), was detected in one index individual but in none of 971 controls. In addition, three sequence alterations were detected in introns (Table 1), but none of them seemed disease-related.

c.1592delT and 3433G→C were then introduced into PALB2-expressing complementary DNA vectors and tested functionally. As shown in Fig. 1a, b, c.1592delT resulted in a truncated protein (PALB2-L531Fs), which had a markedly decreased BRCA2-binding affinity without affecting endogenous BRCA2 abundance upon transient overexpression (Fig. 1b). Consistent with this observation and the functional importance of BRCA2–PALB2 complex formation, PALB2-L531Fs failed to support HR in PALB2-knockdown cells (Fig. 1c) or to restore crosslink repair in PALB2-deficient cells (Fig. 1d). Thus, c.1592delT is a genuine loss-of-function mutation. In contrast, PALB2-G1145R seemed to be fully capable of BRCA2 binding and was functional in these two assays.

Subsequently, *PALB2* c.1592delT was sought in germline DNAs of unselected (not selected for or against family history of cancer) female breast cancer cases ($n = 1{,}918$), unselected male breast cancers ($n = 141$), colorectal cancers ($n = 476$; 188 were familial and 288 belonged to the unselected group) and prostate cancer cases ($n = 639$; 164 were familial and 475 belonged to the unselected group), all from Finland. In the unselected female breast cancer group, 18 additional mutation-positive cases were identified, and a

[1]Department of Clinical Genetics, [2]Department of Surgery, [3]Department of Oncology, [4]Department of Pathology, University of Oulu and Oulu University Hospital, FIN-90029 OYS, Finland. [5]Dana-Farber Cancer Institute and Harvard Medical School, 44 Binney Street, Boston, Massachusetts 02115, USA. [6]Laboratory of Cancer Genetics, Institute of Medical Technology, FIN-33520 University of Tampere and Tampere University Hospital, Finland. [7]Institute of Clinical Medicine, Pathology and Forensic Medicine, University of Kuopio and Department of Clinical Pathology, Kuopio University Hospital, FIN-70211 Kuopio, Finland. [8]Cancer Center and Department of Pathology, Massachusetts General Hospital and Harvard Medical School, Charlestown, Massachusetts 02129, USA. [9]Department of Medical Genetics, University of Helsinki and Biomedicum Helsinki, FIN-00014 HY, Finland. [10]Department of Biosciences and Nutrition, and Clinical Research Centre, Karolinska Institutet, SE-17177 Huddinge, Sweden. [11]Institute of Clinical Medicine, Oncology, University of Kuopio and Department of Oncology, Kuopio University Hospital, FIN-70211 Kuopio, Finland. †Present address: National Human Genome Research Institute/NIH, Bethesda, Maryland 20892, USA. *These authors contributed equally to the work.

Table 1 | *PALB2* sequence variants observed in comprehensive mutation screening in familial breast cancer patients

Exon/intron	Nucleotide change	Effect on protein	Carrier frequency, % (*n*/*N*)		*P* (OR; 95% CI)
			Familial cases	Controls	
Ex 4	1010*T*→C	Leu337Ser	8.0 (9/113)	10.0 (31/304)	0.5 (0.8; 0.4–1.7)
Ex 4	1592delT	Leu531→Fs→Stop	2.7 (3/113)	0.2 (6/2,501)	0.005 (11.3; 1.8–57.8)
Ex 4	1676A→G	Gln559Arg	8.8 (10/113)	15.0 (150/999)	0.08 (0.6; 0.3–1.1)
Ex 8	2794G→A	Val932Met	1.8 (2/113)	5.1 (16/315)	0.1 (0.3; 0.1–1.5)
Ex 12	3300*T*→G	Thr1100Thr	3.5 (4/113)	1.6 (5/322)	0.2 (2.3; 0.6–8.8)
Ex 13	3433G→C	Gly1145Arg	0.9 (1/113)	0 (0/971)	0.1 (n.a.)
IVS1–46	G→A	–	8.8 (10/113)	6.4 (19/295)	0.4 (1.4; 0.6–3.1)
IVS4–70	*T*→G	–	1.8 (2/113)	0 (0/302)	0.07 (n.a.)
IVS4–58	A→C	–	6.1 (7/113)	2.3 (7/302)	0.07 (2.8; 0.95–8.1)

The following sequence information was used: NC_000016.8 (genomic DNA) and NM_024675.2 (mRNA and protein). Fs, frameshift; n.a., not available.

significant association between c.1592delT and female breast cancer was observed (0.9% versus 0.2% in controls ($n = 2{,}051$; 70.6% females and 29.4% males), $P = 0.003$, OR 3.94, 95% CI 1.5–12.1). In addition, the mutant allele was detected in one familial prostate cancer individual, but none of the male breast cancer or colorectal cancer cases revealed any evidence of *PALB2* involvement (Supplementary Table 1). Because all unselected cases and the familial prostate and colorectal cases were only screened for c.1592delT, possible involvement of other, potentially disease-related *PALB2* mutations in these cancers cannot be ruled out. Of a total of 22 identified unrelated cancer patients (21 breast and 1 prostate) heterozygous for *PALB2* c.1592delT, 16 were tested for possible co-segregation of known Finnish *BRCA1* and *BRCA2* mutations[8], and none was detected.

The average age of disease onset for c.1592delT mutation-positive individuals was 52.9 years (variation 39–73 years), which seems slightly younger than the average of the remaining individuals in the unselected breast cancer group (57.8 years, variation 23–95 years; $P = 0.17$) but older than those with Finnish *BRCA1* (46 years, variation 32–57 years) and *BRCA2* (48 years, variation 45–67 years) mutations[9]. However, additional studies will be required to address the influence of the *PALB2* mutation on the age of disease onset. The mutation was also observed in six controls (0.2%; 6/2,501), suggesting that the penetrance of c.1592delT is incomplete. However, most control individuals heterozygous for the mutation were relatively young (five females aged between 27 and 51 years and one male aged 28 years), compared with the above-noted average age of disease onset for affected c.1592delT carriers. The actual penetrance might therefore be higher than currently observed.

For the 18 unselected mutation-positive breast cancer patients, available records were analysed for evidence of a positive family history, and at least half of these families were found to have an apparently heritable disease history (for example, a–d and g in Fig. 2). In addition to breast cancer, all families studied showed other forms of cancer, including colorectal, stomach, endometrial and pancreatic cancers and leukaemia (Fig. 2). Segregation analysis of the truncation allele with regard to cancer incidence was attempted in three of the families with breast cancer studied (Fig. 2a–c) but was not sufficiently informative to draw meaningful conclusions because of a lack of DNA samples from suitable family members. For the remaining families, the analysis was restricted only to the affected index individual who initially displayed the c.1592delT allele. A segregation

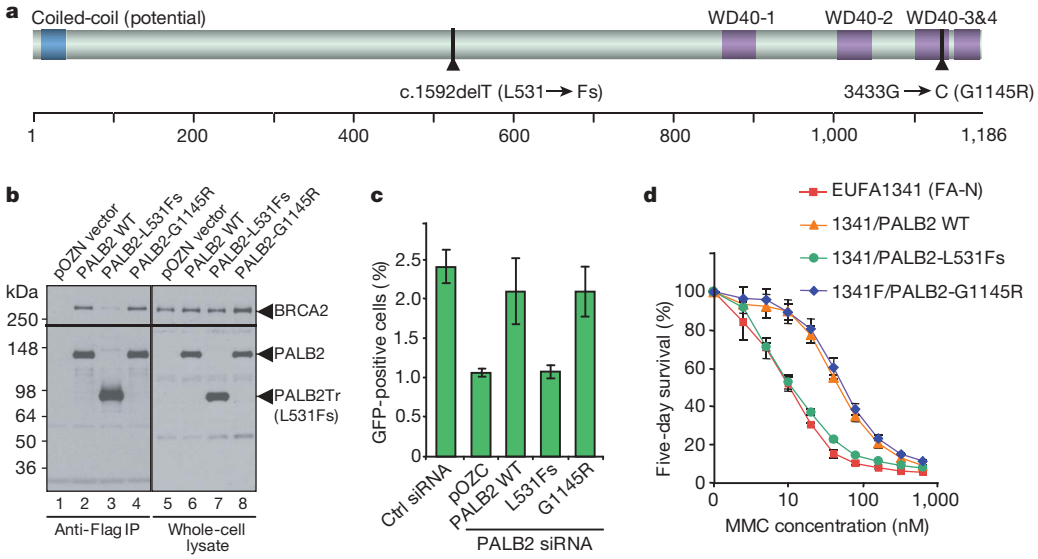


Figure 1 | PALB2 protein structure and assessment of functional consequences of the c.1592delT and 3433G→C sequence alterations. **a**, Schematic diagram of the protein showing predicted functional domains and the sites of the two main sequence changes observed. **b**, 293T cells were transfected with the indicated plasmids and PALB2 proteins double-tagged with Flag and haemagglutinin (HA) were immunoprecipitated (IP) with anti-Flag M2 agarose beads. The abundance of tagged PALB2 proteins and BRCA2 in the precipitates was analysed by western blotting (lanes 1–4). The endogenous BRCA2 abundance and relative levels of ectopic PALB2 expression are shown on the right (lanes 5–8). WT, wild type. The numbers at the left indicate the positions of molecular mass markers. **c**, DR-U2OS HR reporter cells were treated with control or PALB2 short interfering RNAs (siRNAs) and then co-transfected with pCBASce together with the pOZC vector or cDNA constructs. Cells positive for green fluorescent protein were counted 72 h later. The PALB2 cDNAs contain seven silent base changes and are resistant to the PALB2 siRNA. The results shown are means ± s.d. for three independent experiments, each performed in duplicate. **d**, EUFA1341 (FA-N) fibroblasts stably expressing indicated PALB2 proteins were treated with the indicated concentrations of MMC and their survival was assayed 120 h after treatment. The results are means ± s.d. from a representative experiment performed in triplicate.

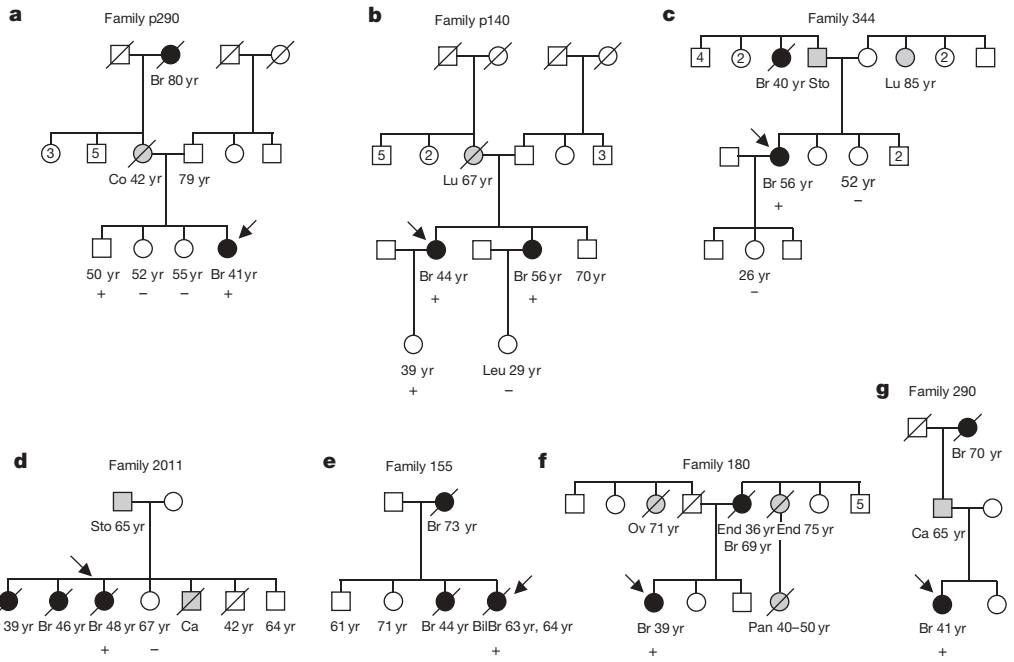


Figure 2 | Examples of pedigrees of breast cancer families exhibiting the c.1592delT allele. a–c, Families attempted for partial segregation analysis. **d–g**, Families displaying the mutant allele but in which segregation analysis was not possible because of a lack of DNA samples from suitable family members. Patients with breast cancer are marked with black circles (BilBr, bilateral breast cancer). Other cancer types are marked in grey and mentioned when known (Ca, cancer of unknown type with liver metastases (the patient died at the age of 68 years); Co, colorectal; End, endometrial; Lu, lung; Leu, leukaemia; Ov, ovarian; Pan, pancreatic; Sto, stomach). Age at cancer diagnosis is shown, when known. Individuals genotyped for c.1592delT are marked with either a plus sign (if mutation-positive) or a minus sign (if mutation-negative). Age at monitoring is shown for healthy individuals genotyped for c.1592delT. Index individuals are marked with arrows. A slashed pedigree symbol indicates a deceased individual.

study was also performed on the family of the mutation-positive patient with prostate cancer (Fig. 3). Other than the individual who died early at 52 years of age, all male carriers developed prostate cancer by the age of 76 years, indicating high penetrance of the mutation in the two generations of this family that were studied.

To test whether loss of heterozygosity (deletion of the wild-type allele) had occurred in tumours of individuals heterozygous for the mutation, genomic DNA extracted from formalin-fixed, paraffin-embedded tumour sections from six patients was analysed. Before DNA extraction, laser-capture microdissection was performed to isolate pure cancer cell populations from tumours from five patients. *PALB2* gene segments (about 200 base pairs) surrounding the c.1592delT mutation were amplified by polymerase chain reaction with multiple primer pairs, and sequenced. The existence of the c.1592delT mutation was confirmed in all tumours, whereas no reproducible evidence of loss of heterozygosity was ever observed, implying that these tumours were likely to have been driven, at least in part, by PALB2 haploinsufficiency, perhaps in combination with a dominant-negative effect of the truncated protein product.

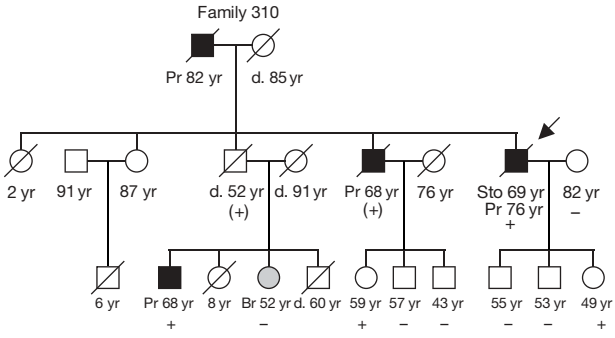


Figure 3 | Pedigree of a prostate cancer family segregating the *PALB2* c.1592delT truncation allele. Individuals with prostate cancer are marked with black squares. In addition to prostate cancer, the index individual (marked with an arrow) also had stomach cancer (Sto). The single individual with breast cancer (Br) is shown in grey. No other individuals with cancer were known to have occurred in this family. Individuals genotyped for c.1592delT are marked with either a plus sign (if mutation positive) or a minus sign (if mutation negative). Age at monitoring is shown for healthy individuals typed for c.1592delT. A slashed symbol indicates a deceased individual. A plus sign in parentheses indicates an obligate mutation carrier. It was not possible to obtain analysable DNA from the first-generation male ancestor.

Immunohistochemistry was also performed on sections from the same six tumours noted above and from one further sample. As shown in Supplementary Table 2, all except one revealed strong expression of oestrogen receptor, and five of seven showed expression of progesterone receptor. These results imply that *PALB2* tumours share the above phenotypic properties with those generated by *BRCA2* mutations[10,11]. However, because of the limited number of specimens analysed, more extensive analysis is merited.

Very recent discoveries have revealed that, like BRCA2, PALB2 is also a Fanconi anaemia gene product, and *PALB2* and *BRCA2* patients with Fanconi anaemia share similar severe phenotypes characterized by childhood cancers and early mortality[12–14]. Thus, in addition to their role in the development of Fanconi anaemia and childhood cancers, the present results indicate that suitable germline mutations in *PALB2* also confer an elevated risk of breast cancer and perhaps prostate cancer.

PALB2 c.1592delT, which subsequently proved to be a founder mutation (data not shown), seems to be associated with a roughly fourfold increased hereditary propensity for female breast cancer and to make a limited contribution to familial prostate cancer. The truncated protein product seems to be stable but is functionally defective, being unable to support intact BRCA2 DNA repair function; this observation, together with the above-noted significant statistical difference in its prevalence between patients with breast cancer and controls, indicates that this mutation is a significant component of heritable susceptibility to breast cancer in Finland. The roughly 1%

occurrence rate of c.1592delT in unselected breast cancer patients is remarkable, especially because the 19 different pathogenic mutations identified in *BRCA1* and *BRCA2* together account for about 1.8% of Finnish breast cancer cases[8]. In Finland, about 4,000 women are diagnosed with breast cancer annually. This single mutation could therefore be responsible for about 40 new cases per year. Furthermore, because the *PALB2* gene was comprehensively screened in only 113 cancer families, the existence of other cancer-predisposing *PALB2* alterations in the Finnish population remains possible. The present results also imply that *PALB2* might be a significant new cancer susceptibility gene in other populations. In keeping with this notion, two of the mutations identified in Fanconi anaemia patients in non-Finnish populations seem to be associated with incidences of familial breast cancer[12,13].

METHODS

Sample selection. See Supplementary Information for details.

Mutation screening. The entire coding region and exon–intron boundaries of the *PALB2* gene were screened for germline mutations by conformation-sensitive gel electrophoresis[15,16] and direct sequencing.

Statistical and bioinformatic methods. Carrier frequencies were compared using Pearson's χ^2 test or Fisher's exact test. SPSS version 12.0 for Windows was used. The Mann–Whitney *U*-test was used to compare mean ages of disease onset between mutation carriers and non-carriers. PolyPhen simulation, ESEfinder 2.0 and NNSplice software were used to predict functions of the observed sequence changes. The cutoff value for statistical significance was $P = 0.01$.

Functional analysis. The retroviral PALB2 cDNA vectors, pOZN-PALB2 and pOZC-PALB2, have been described previously[4]. The mutations, c.1592delT and 3433G→C, were introduced into these vectors by site-directed mutagenesis with the QuikChange method (Stratagene). The HR/DSBR assay was performed as described[4]. The generation of EUA1341(FA-N) fibroblasts stably expressing various PALB2 species and subsequent mitomycin-C (MMC) sensitivity assays were as described[12].

Received 5 December 2006; accepted 19 January 2007.
Published online 7 February 2007; corrected 16 February 2007 (details online).

1. Wooster, R. & Weber, B. L. Breast and ovarian cancer. *N. Engl. J. Med.* **348,** 2339–2347 (2003).
2. Seal, S. *et al.* Truncating mutations in the Fanconi anemia J gene *BRIP1* are low-penetrance breast cancer susceptibility alleles. *Nature Genet.* **38,** 1239–1241 (2006).
3. Renwick, A. *et al.* ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. *Nature Genet.* **38,** 873–875 (2006).
4. Xia, B. *et al.* Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2. *Mol. Cell* **22,** 719–729 (2006).
5. Shivji, M. K. & Venkitaraman, A. R. DNA recombination, chromosomal stability and carcinogenesis: insights into the role of BRCA2. *DNA Repair (Amst.)* **3,** 835–843 (2004).
6. King, M. C., Marks, J. H. & Mandell, J. B. Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2. *Science* **302,** 643–646 (2003).
7. Howlett, N. G. *et al.* Biallelic inactivation of BRCA2 in Fanconi anemia. *Science* **297,** 606–609 (2002).
8. Syrjäkoski, K. *et al.* Population-based study of BRCA1 and BRCA2 mutations in 1035 unselected Finnish breast cancer patients. *J. Natl Cancer Inst.* **92,** 1529–1531 (2000).
9. Sarantaus, L. *et al.* Multiple founder effects and geographical clustering of BRCA1 and BRCA2 families in Finland. *Eur. J. Hum. Genet.* **8,** 757–763 (2000).
10. Borg, Å. Molecular and pathological characterization of inherited breast cancer. *Semin. Cancer Biol.* **11,** 375–385 (2001).
11. Hedenfalk, I. *et al.* Molecular classification of familial non-BRCA1/BRCA2 breast cancer. *Proc. Natl Acad. Sci. USA* **100,** 2532–2537 (2003).
12. Xia, B. *et al.* Fanconi anemia is associated with a defect in the BRCA2 partner PALB2. *Nature Genet.* advance online publication, doi:10.1038/ng1942 (31 December 2006).
13. Reid, S. *et al.* Biallelic mutations in *PALB2* cause Fanconi anemia subtype FA-N and predispose to childhood cancer. *Nature Genet.* advance online publication, doi:10.1038/ng1947 (31 December 2006).
14. Alter, B. P., Rosenberg, P. S. & Brody, L. C. Clinical and molecular features associated with biallelic mutations in FANCD1/BRCA2. *J. Med. Genet.* **44,** 1–9 (2007).
15. Ganguly, A., Rock, M. J. & Prockop, D. J. Conformation-sensitive gel electrophoresis for rapid detection of single-base differences in double-stranded PCR products and DNA fragments: evidence for solvent-induced bends in DNA heteroduplexes. *Proc. Natl Acad. Sci. USA* **90,** 10325–10329 (1993).
16. Körkkö, J., Annunen, S., Pihlajamaa, T., Prockop, D. J. & Ala-Kokko, L. Conformation sensitive gel electrophoresis for simple and accurate detection of mutations: comparison with denaturing gradient gel electrophoresis and nucleotide sequencing. *Proc. Natl Acad. Sci. USA* **95,** 1681–1685 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Ignatius, E. Nieminen, K. Mononen, H. Konola, O. Kajula, M. Vahera, K. Holli, T. Tammela, K. Rouhento, L. Enroth, R. Vaalavuo and S. Marttinen for help in sample and data collection and technical assistance. We also thank the Finnish Red Cross Blood Service for help with collection of population control blood samples, the Finnish Cancer Registry for information on cancer occurrence, and all patients and their family members for volunteering to participate in these studies. This study was supported by the Foundation for the Finnish Cancer Institute, the Academy of Finland, the Ida Montin Foundation, the Cancer Foundation of Northern Finland, the University of Oulu, Oulu University Hospital, the Reino Lahtikari Foundation, the Sigrid Juselius Foundation, Competitive Research Funding of the Pirkanmaa Hospital District, and grants to D.M.L. from the National Cancer Institute. This work was also supported by a grant from the Shapiro Family Foundation. D.M.L. is a scientific consultant to and a grant recipient of The Novartis Institute for Biomedical Research.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to R.W. (robert.winqvist@oulu.fi) or D.M.L. (david_livingston@dfci.harvard.edu).

Vol 446|15 March 2007|doi:10.1038/nature05585

LETTERS

A Hedgehog- and Antennapedia-dependent niche maintains *Drosophila* haematopoietic precursors

Lolitika Mandal[1], Julian A. Martinez-Agosto[3], Cory J. Evans[1], Volker Hartenstein[1] & Utpal Banerjee[1,2]

The *Drosophila melanogaster* lymph gland is a haematopoietic organ[1–3] in which pluripotent blood cell progenitors proliferate and mature into differentiated haemocytes. Previous work[4] has defined three domains, the medullary zone, the cortical zone and the posterior signalling centre (PSC), within the developing third-instar lymph gland. The medullary zone is populated by a core of undifferentiated, slowly cycling progenitor cells, whereas mature haemocytes comprising plasmatocytes, crystal cells and lamellocytes are peripherally located in the cortical zone. The PSC comprises a third region that was first defined as a small group of cells expressing the Notch ligand Serrate[5]. Here we show that the PSC is specified early in the embryo by the homeotic gene *Antennapedia* (*Antp*) and expresses the signalling molecule Hedgehog. In the absence of the PSC or the Hedgehog signal, the precursor population of the medullary zone is lost because cells differentiate prematurely. We conclude that the PSC functions as a haematopoietic niche that is essential for the maintenance of blood cell precursors in *Drosophila*. Identification of this system allows the opportunity for genetic manipulation and direct *in vivo* imaging of a haematopoietic niche interacting with blood precursors.

The *Drosophila* lymph gland primordium is formed by the coalescence of three paired clusters of cells that express Odd-skipped (Odd) and arise within segments T1–T3 (Fig. 1a) of the embryonic cardiogenic mesoderm[6]. At developmental stages 11–12, mesodermal expression of Antp is restricted to the T3 segment (Fig. 1b, c). A fraction of these Antp-expressing cells will contribute to the formation of the dorsal vessel[7,8], whereas the remainder, which also express

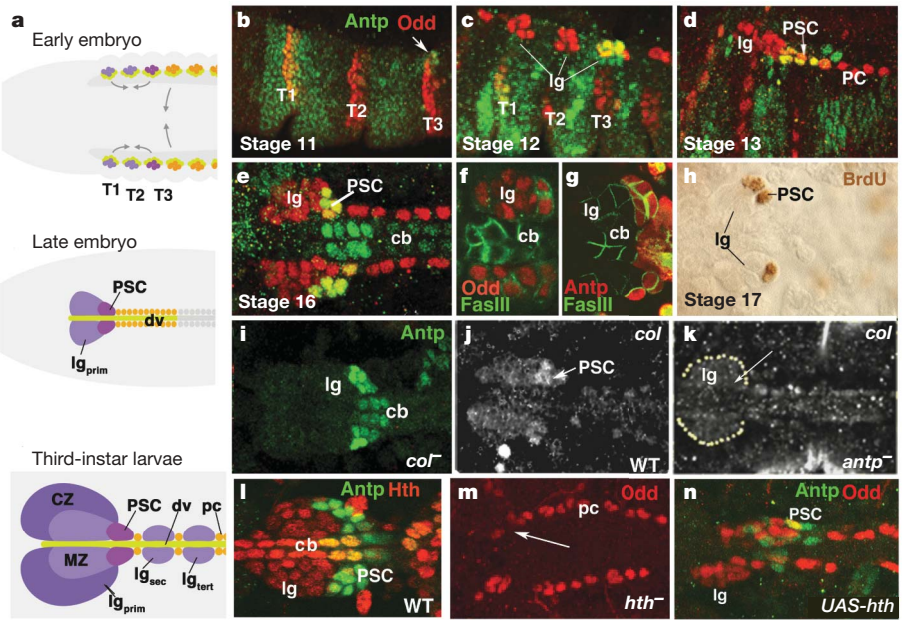
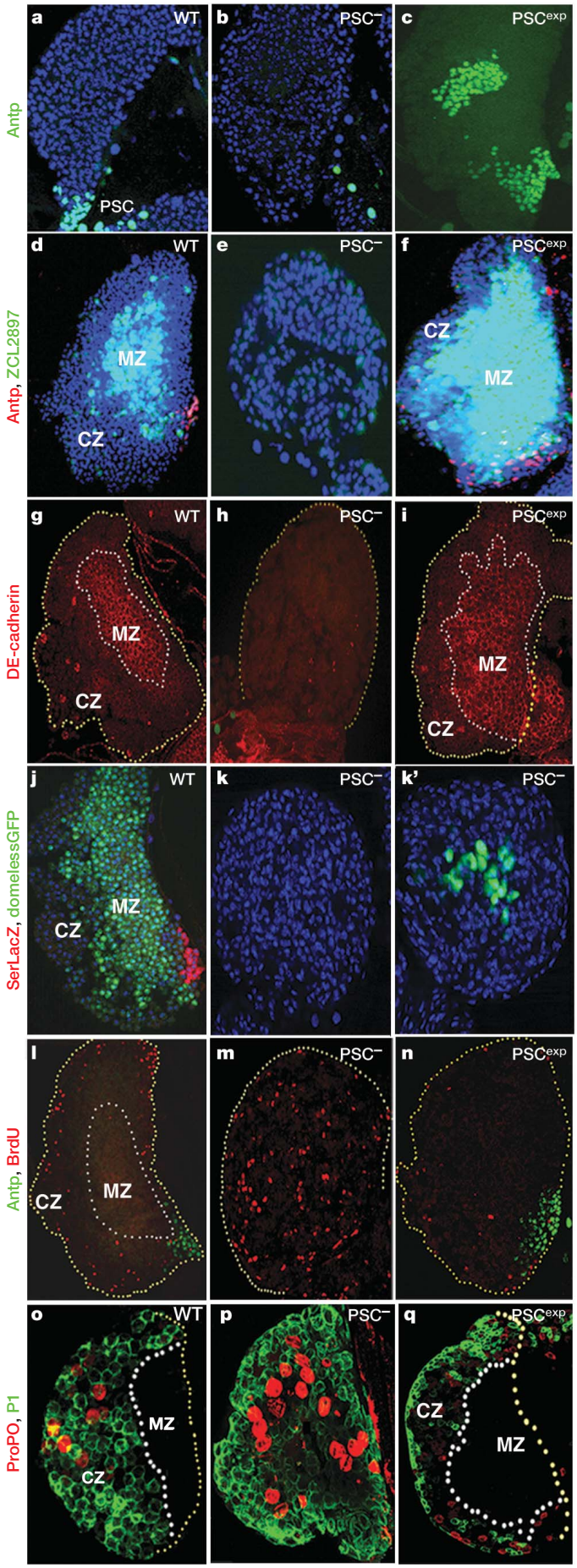


Figure 1 | Embryonic specification of the PSC by Antp. **a**, Schematic representation of the development of the *Drosophila* lymph gland. T1–T3, three thoracic segments; dv, the dorsal vessel; cb, cardioblast; lg_{prim}, lg_{sec} lg_{tert}, primary, secondary and tertiary lobes, respectively, of the lymph gland; and pc, pericardial cells. **b**–**e**, Immunohistochemical analysis showed that the lymph gland is formed by the fusion of the three Odd-positive cell groups (T1, T2 and T3). Antp expression is confined to a group of cells at the posterior boundary of T3 (**b**, **e**). The cells that will become larval PSC remain confined to the posterior edge of the embryonic lymph gland (**e**). **f**, **g**, Fas III (Fasciclin III, a homophilic cell adhesion molecule) is upregulated in the PSC. **h**, In the late embryo, the PSC cells incorporate BrdU. **i**, Antp expression is maintained in the embryonic *col* mutant background. **j**, **k**, Expression of *col*, detected by *in situ* hybridization, in the PSC (WT, arrow in **j**) is eliminated in an *Antp* mutant background (arrow in **k**). **l**, Antp protein is expressed in the PSC, whereas Hth protein is seen in the rest of the lymph gland. **m**, **n**, In the *hth* mutant background (**m**), the lymph gland is virtually eliminated (arrow). Overexpression of *hth* (*twist-gal4, UAS-hth*; **n**) causes a reduction in the number of cells in the PSC. All colours correspond to the marker label in each panel. All images were acquired using a ×40 objective with additional ×1.3 (**b**, **c**, **d**) or ×2.5 (**e**–**n**) confocal magnification.

[1]Department of Molecular, Cell and Developmental Biology, Molecular Biology Institute, [2]Department of Biological Chemistry, [3]Department of Pediatrics, Mattel Children's Hospital at UCLA, University of California, Los Angeles, Los Angeles, California 90095, USA.



Odd, give rise to the PSC (Fig. 1d, e). By stages 13–16, the clusters coalesce and Antp is observed in 5–6 cells at the posterior boundary of the lymph gland (Fig. 1d, e). The expression of Antp is subsequently maintained in the PSC through the third larval instar (see Fig. 2a). The embryonic stage 16 PSC can also be distinguished by Fasciclin III expression (Fig. 1f, g) and at stage 17 these are the only cells in the lymph gland that incorporate BrdU (Fig. 1h).

Previous studies have identified the transcription factor Collier (Col) as an essential component regulating PSC function[9]. The gene for this protein is initially expressed in the entire embryonic lymph gland anlagen and by stage 16 is refined to the PSC. In *col* mutants, the PSC is initially specified, but is entirely lost by the third larval instar. To address further the role of Antp and Col in embryonic lymph gland development, we investigated the expression of each gene in the loss-of-function mutant background of the other. We found that loss of *col* does not affect embryonic Antp expression (Fig. 1i). In contrast, *col* expression is absent in the PSC of *Antp* mutant embryos (Fig. 1j, k), establishing that Antp functions genetically upstream of Col in the PSC.

In imaginal discs, the expression of Antp is related to that of the homeodomain cofactor Homothorax (Hth)[10]. In the embryonic lymph gland, Hth is initially expressed ubiquitously but is subsequently downregulated in PSC cells, which become Antp-positive (Fig. 1l). In *hth* loss-of-function mutants, the lymph gland is largely missing (Fig. 1m), whereas misexpression of *hth* causes loss of PSC and the size of the embryonic lymph gland remains relatively normal (Fig. 1n). We conclude that a mutually exclusive functional relationship exists between Antp and Hth in the lymph gland such that Antp specifies the PSC, whereas Hth specifies the rest of the lymph gland tissue. Interestingly, knocking out the mouse homologue of Hth, Meis1, eliminates definitive haematopoiesis[11,12]. Meis1 is also required for the leukaemic transformation of myeloid precursors overexpressing HoxB9[13].

Although lymph gland development is initiated in the embryo, the establishment of zones and the majority of haemocyte maturation takes place in the third larval instar. At this stage, Antp continues to be expressed in the wild-type PSC (Fig. 2a). To investigate how the loss of PSC cells affects haematopoiesis, we first examined Antp expression in third instar *col* mutant lymph glands. In this background, all Antp-positive PSC cells are missing (Fig. 2b), consistent with the previously described role for *col* in PSC maintenance[9]. Overexpression of Antp within the PSC increases the size of PSC from the usual 30–45 cells to 100–200 cells (Fig. 2c). These PSC cells are scattered over a larger volume, often forming two or three large cell clusters rather than the single, dense population seen in wild type.

Figure 2 | The larval PSC functions as a haematopoietic niche to maintain blood precursors in the medullary zone. **a**, In wild-type (WT), the PSC consists of a cell cluster located along the posterior edge of the lymph gland. **b**, Antp expression in the lymph gland is missing in *col*$^{-}$ (PSC^{-}). The residual expression is in the cardioblasts of the dorsal vessel. **c**, The PSC is greatly expanded on *Antp* overexpression (PSC^{exp}). **d**–**k'**, In wild type, medullary zone markers (**d**, **g**, **j**) are restricted to the precursor population of cells. These markers are all eliminated in the absence of the PSC (**e**, **h**, **k**). Rarely, a much reduced medullary zone can be seen (example shown in **k'**). Expansion of the PSC causes expansion of the medullary zone (**f**, **i**). **l**–**n**, In wild-type (**l**), BrdU incorporation is limited to the cortical zone. In the absence of PSC, cells in more medial regions incorporate BrdU (**m**), whereas in an expanded PSC genotype (**n**), the number of cells in S-phase is greatly reduced. **o**–**q**, In wild-type (**o**), differentiating cells reside in the peripheral cortical zone. In the absence of a PSC (**p**), differentiated cells are found throughout the lymph gland lobe. On expansion of the PSC (**q**), these cells are restricted to a thin layer along the distal edge. The colour of molecular markers for each row corresponds to those of the side labels (the label 'domelessGFP' corresponds to the expression of GFP under the control of *domeless-gal4*; 'SerLacZ' is expression of β-galactosidase under the control of the *Ser9.5* (ref. 4) enhancer). All images were acquired using a 40× objective. The PSC panels (**b**, **e**, **h**, **k**, **k'**, **m**, **p**) also reflect an additional ×1.5 confocal magnification.

To determine the role of PSC in haematopoiesis, we investigated the expression pattern of various markers in lymph glands of larvae of the above genotypes, which either lack a PSC or have an enlarged PSC. The status of blood cell progenitors was directly assessed using the medullary-zone-specific markers[4] *ZCL2897*, DE-cadherin (Shotgun) and *domeless-gal4* (Fig. 2d–k'). In *col* mutant lymph glands, expression of these markers is absent or severely reduced (Fig. 2e, h, k, k') and when the PSC is expanded, the medullary zone is greatly enlarged (Fig. 2f, i). Our previous work demonstrated that medullary zone precursors are relatively quiescent[4], a characteristic similar to the slowly cycling stem cell or progenitor populations in other systems[14]. BrdU incorporation in the wild-type lymph gland is largely restricted to the cortical zone[4] (Fig. 2l), but in third-instar *col* mutants incorporation of BrdU is increased relative to wild type and becomes distributed throughout the lymph gland (Fig. 2m), suggesting that the quiescence of the medullary zone haematopoietic precursors is no longer maintained in the absence of the PSC. Similarly, when the PSC domain is expanded, BrdU incorporation is significantly suppressed throughout the lymph gland (Fig. 2n).

We next used P1 and ProPO as markers for plasmatocytes and crystal cells, respectively, to assess the extent of haemocyte differentiation within lymph glands of the above genotypes. Loss of the PSC does not compromise haemocyte differentiation; rather, mature plasmatocytes and crystal cells are found abundantly within the lymph gland. Furthermore, the distribution of these differentiating cells is not restricted to the peripheral region that normally constitutes the cortical zone and many cells expressing ProPO and P1 can be observed medially throughout the region normally occupied by the medullary zone (Fig. 2o, p). Increasing the PSC domain causes a concomitant reduction in the differentiation of haemocytes (Fig. 2q).

In summary, loss of the PSC causes a loss of medullary zone markers, a loss of the quiescence normally observed in the wild-type precursor population and an increase in cellular differentiation throughout the lymph gland. Similarly, increased PSC size leads to an increase in the medullary zone, a decrease in BrdU incorporation and a decrease in the expression of maturation markers. We conclude that the PSC functions as a haematopoietic niche that maintains the population of multipotent blood cell progenitors within the lymph gland. The observed abundance of mature cells in the absence of the PSC suggests that the early blood cell precursors generated during the normal course of development will differentiate in the absence of a PSC-dependent mechanism that normally maintains progenitors as a population. This situation is reminiscent of the *Drosophila*[15] and *Caenorhabditis elegans*[16] germ lines in which disruption of the niche does not block differentiation *per se*, but lesser numbers of differentiated cells are generated as a result of the failure to maintain stem cells. It is also interesting to note that *col* mutant larvae are unable to mount a lamellocyte response to immune challenge[9]. We speculate that this could be because of the loss of precursor cells that are necessary as a reserve to differentiate during infestation.

Recent work on several vertebrate and invertebrate developmental systems has highlighted the importance of niches[14,17] as unique microenvironments in the maintenance of precursor cell populations. Examples include haematopoietic[18,19], germline[20] and epidermal[21] stem cell niches that provide, through complex signalling interactions, stem cells with the ability to self-renew and persist in a non-differentiated state. The work presented in this report demonstrates that the PSC is required for the maintenance of medullary zone haematopoietic progenitors. The medullary zone represents a group of cells within the lymph gland that are compactly arranged and express the homotypic cell-adhesion molecule, DE-cadherin[4]. These cells are pluripotent, slowly cycling and undifferentiated and are capable of self-renewal. It is presently uncertain whether *Drosophila* has blood stem cells capable of long-term repopulation as haematopoietic stem cells are in vertebrates. Nevertheless, it is clear that the maintenance of medullary zone cells as precursors is niche dependent.

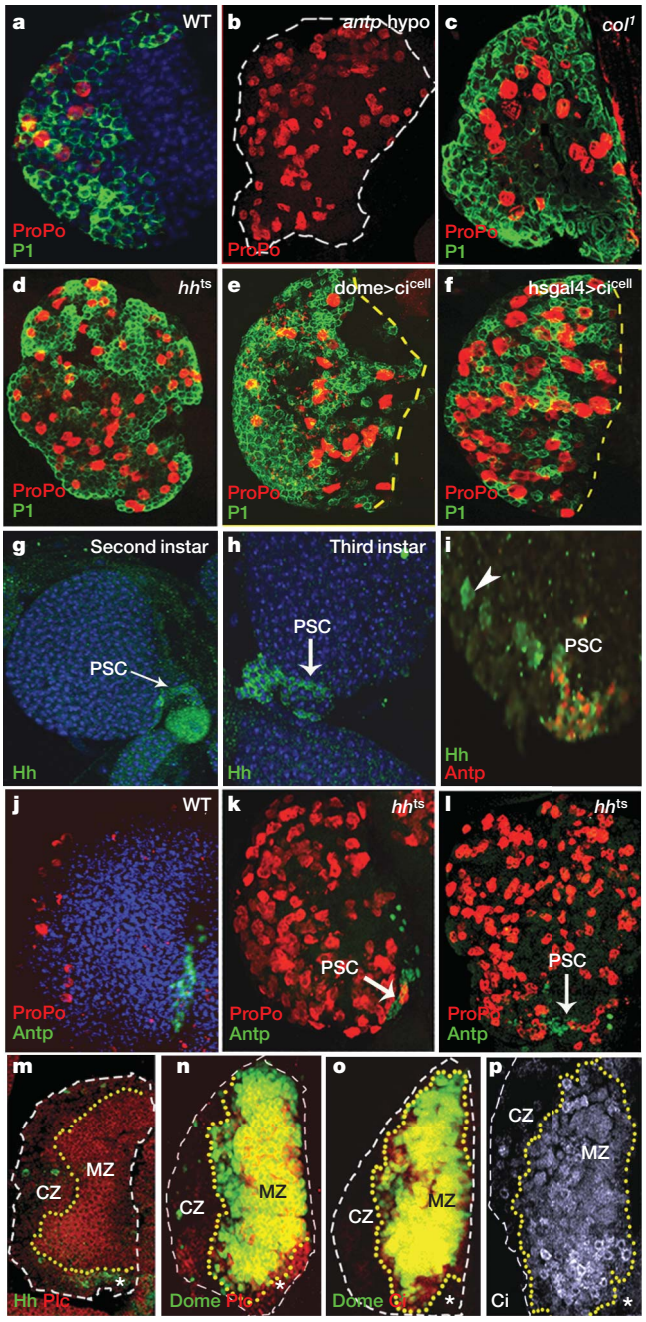


Figure 3 | A Hedgehog signal from the PSC is required for the maintenance of the precursor cell population of the medullary zone. **a**, Wild-type (WT) expression of the cortical zone markers P1 (green) and ProPO (red). **b**, Third-instar *Antp*-mutant[22] larvae show a lymph gland phenotype similar to that seen in *col* mutants (**c**). hh^{ts2} (**d**) shows an identical phenotype (compare to **c**). Misexpression of dominant-negative Ci (**e**, **f**) phenocopies *Antp*, *col* and *hh* loss-of-function lymph gland phenotypes (**b**–**d**). **g**–**i**, Hh expression (green) is restricted to the PSC in the second (**g**) and third instars (**h**, **i**). All cells expressing Antp (red in **i**) also express Hh (yellow in **i**). A few dispersed cells (arrowhead) in the cortical zone also initiate Hh expression (**i**). **j**–**l**, The PSC (Antp in green) is present in hh^{ts2} mutants (**k**, **l**) as in wild-type (**j**). **m**, Hh expression marks the PSC and Ptc marks the medullary zone. **n**–**p**, Ptc and Ci (red) expression co-localizes with the medullary zone marker *domeless-gal4, UAS-GFP* (labelled 'Dome'). Ci forms a gradient with highest staining intensity near the PSC. The asterisk marks the PSC. Haematopoietic markers for each panel are colour-coded. All images were acquired using a ×40 objective. Panels (**c**, **d**, **k** and **l**) reflect an additional ×1.5 confocal magnification. The confocal magnification used for the close-up of the PSC in panel (**i**) was ×3.

In order for the PSC to function as a haematopoietic niche there should exist a means by which the PSC can communicate with precursors. As such, a signal emanating from the PSC and sensed by the medullary zone represents an attractive model of how this might occur. Although we have reported that Ser (ref. 5) and Upd3 (ref. 4) are expressed in the PSC, preliminary analysis suggests that elimination of either of these ligands alone will not cause the phenotype seen for *Antp* and *col* mutants. We therefore investigated the haematopoietic role of several signalling pathways and identified the *hedgehog* (*hh*) signalling pathway as a putative regulator in the maintenance of blood cell progenitors. The hh^{ts2} lymph gland (Fig. 3d) is remarkably similar in its phenotype to that seen for *Antp* hypomorphic[22] (Fig. 3b) or *col* loss-of-function (Fig. 3c) mutants (compare to wild type in Fig. 3a). Blocking Hh signalling in the lymph gland through the expression of a dominant-negative form of the downstream activator Cubitus interruptus (Ci, the *Drosophila* homologue of Gli) also causes a phenotype similar to that observed in *Antp* and *col* loss-of-function backgrounds. This is true when expressed either specifically in the medullary zone (Fig. 3e) or throughout the lymph gland (Fig. 3f).

Consistent with the above functional results, Hh protein is expressed in the second instar PSC (Fig. 3g) and continues to be expressed in third instar PSC cells (Fig. 3h, i). In the hh^{ts2} mutant background, the PSC cells continue to express Antp at the restrictive temperature (Fig. 3j–l) indicating that, unlike *col* and *Antp*, Hh is not essential for the specification of the PSC. Rather, Hh constitutes a component of the signalling network that allows the PSC to maintain the precursor population of the medullary zone. Consistent with this notion, downstream components of the Hh pathway, the receptor Patched (Ptc) (Fig. 3m, n) and activated Ci (Fig. 3o, p), are found in the medullary zone. On the basis of both functional and expression data, we propose that Hh in the PSC signals through activated Ci in medullary zone cells, thereby keeping them in a quiescent precursor state.

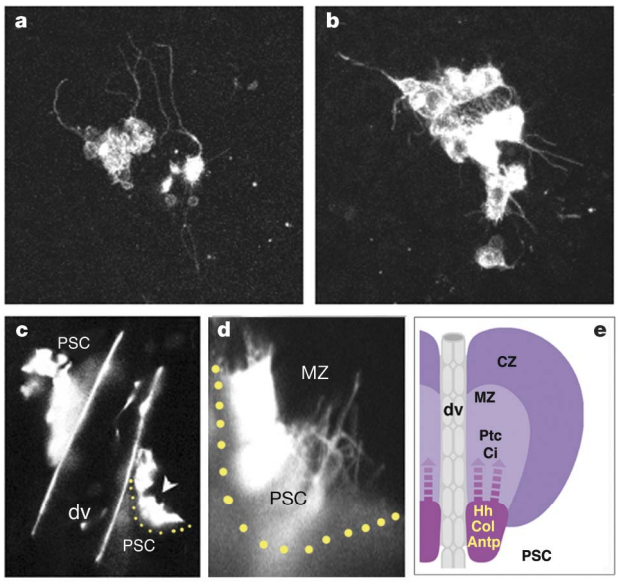


Figure 4 | PSC cells exhibit extensive processes that project into the lymph gland. a, **b**, GFP expressed exclusively in PSC cells (using *Antp-gal4, UAS-mCD8–GFP*) reveals the presence of numerous thin processes that extend over several cell diameters into the medullary zone. **c**, **d**, Analysis of PSC morphology in live animals. Whole-mount live third instar larva showing GFP expression in the PSC. A close-up view (**d**) of the region indicated by the arrowhead in (**c**), showing processes extending into the lymph gland. dv, dorsal vessel. **e**, A schematic representation of the PSC region of the lymph gland as a niche involved in the maintenance of medullary zone progenitors. All images were acquired using a ×40 objective except panel **c**, which was taken with a ×10 objective. Panels **a**,**b** and **c** and **d**, reflect additional ×6, ×2.5, and ×8 confocal magnification, respectively.

The Hh pathway has been studied extensively in the context of animal development[23]. Although the Hh signal does not disperse widely on secretion, many studies have shown that this signal can be transmitted over long distances[24]. The mechanism by which this occurs is not fully clear and this is also true of how the PSC delivers Hh to medullary zone progenitors. However, when labelled with green fluorescent protein (GFP), we find that PSC cells extend numerous thin processes over many cell diameters (Fig. 4). The morphology of the PSC cells, taken together with the long-range function of Hh revealed by the mutant phenotype, indicates that the long cellular extensions may deliver Hh to receiving cells not immediately adjacent to the PSC. In this respect, the *Drosophila* haematopoietic system shows remarkable similarity to the *C. elegans* germline[25]. In both cases, precursors are maintained as a population over some distance from the niche and in both instances, the niche cells extend long processes when interacting with the precursors.

Several studies have highlighted the importance of homeodomain proteins in stem cell development and leukaemias[26,27]. Likewise, the role of Hh in vertebrate and invertebrate stem cell maintenance has recently received much attention[28–30]. The work presented here describes direct roles for Antp in the specification and Hh in the functioning of a haematopoietic niche. The medullary zone cells are blood progenitors that are maintained in the lymph gland at later larval stages by Hh, a signal that originates in the PSC (Fig. 4e). The maintenance of these progenitors provides the ability to respond to additional developmental or immune-based haematopoietic signals. On the basis of these findings, understanding the specific roles of Hh signalling and Hox genes in the establishment and function of vertebrate haematopoietic niches warrants further investigation. The identification of a haematopoietic niche in *Drosophila* will allow future investigation of *in vivo* niche/precursor interactions in a haematopoietic system that allows direct observation, histological studies and extensive genetic analysis.

METHODS

Fly stocks and crosses. The following *Drosophila* strains (donors in parentheses) were used in the described experiments: *Antp-gal4/TM3, Sb* (Cohen, S.M.), GFP-trap line *ZCL2897* (Cooley, L.), hh^{ts2} (Moses, K.), *UAS-*Ci^{cell} (Basler, K.), col^{1}/*CyO, twist-lacZ* and col^{1}; *P(col5-cDNA)/CyO-TM6B, Tb* (Crozatier, M.), *domeless-gal4 (PG14;* Noselli, S.*)* and hth^{64-1}/*TM6B* and *UAS-hth* (Salzberg, A.). The following stocks were obtained from the Bloomington Stock Centre: *UAS-Antp*; *Df(2R)knSA3*; $Antp^{25}$ red^{1} e^{1}/*TM3, Sb*; $Antp^{17}$/*TM3, Sb*; *UAS-2xEYFP*; *hs-gal4/CyO* and *w; P{tubP-gal80[ts]}10.* In *collier* loss-of-function experiments, *domeless-gal4, UAS-2xEYFP* (X chromosome) was placed in the background of heterozygotes carrying the col^{1} allele over the deficiency *Df(2R)knSA3.* The *ZCL2897* intron trap line (X chromosome) was crossed into the background of col^{1} allele homozygotes. In overexpression studies, *UAS-Antp* was driven by *Antp-gal4* and *ZCL2897* was placed in this background to visualize the status of the medullary zone. Similarly, *UAS-hth* was driven by *twist-GAL4* in *hth* overexpression studies. For *Antp* loss-of-function, a heteroallelic combination ($Antp^{25}$/$Antp^{17}$)[22] that survives to the late third instar was used. For misexpression of *UAS-*Ci^{cell}, these flies were crossed to *domeless-gal4; P{tubP-gal80[ts]}10*, grown at 18 °C until hatching, and then transferred to 29 °C until dissection. Additionally, *UAS-*Ci^{cell} was crossed to *hs-gal4*, grown at 18 °C until mid-second instar, heat shocked at 37 °C for 2 h, and then returned to 18 °C until dissection. For hh^{ts2} experiments, the full phenotype is observed when the larvae are shifted to the non-permissive temperature in the mid-second instar.

Immunohistochemistry. Embryos and lymph glands were stained as previously described[5,7]. The following antibodies were used: rabbit anti-Hth (Salzberg, A.), rabbit anti-Hh (Ingham, P.), mouse anti-Antp, mouse anti-Ptc and mouse anti-FasIII (Developmental Studies Hybridoma Bank), rat anti-ProPO antibody (Müller, H.), mouse anti-P1 (Ando, I.), rat anti-DE-cadherin (Hartenstein, V.) and 2A12 rat anti-Ci (Holmgren, R.). Samples were imaged using a BioRad Radiance 2000 confocal with LaserSharp 2000 acquisition software.

Imaging PSC. For imaging the PSC using membrane GFP (*Antp-gal4, UAS-mCD8-GFP*), lymph glands were dissected in ice-cold 1× PBS and fixed in 4% formaldehyde/1× PBS on ice for five minutes. Lymph glands were then placed on a slide in 1× PBS, a cover slip was placed and the tissue was immediately imaged using standard confocal microscopy techniques. For live imaging of the PSC, individual whole *Antp-gal4, UAS-GFP* larvae were washed in water, placed

on a slide dorsal-side up in glycerol, and then a cover slip was placed with sufficient pressure to immobilize the larva long enough for the acquisition of confocal sections through the PSC.

Received 21 September 2006; accepted 8 January 2007.

1. Rizki, T. M. The circulatory system and associated cells and tissues. In *The Genetics and Biology of Drosophila* (eds Ashburner, M. and Wright, T. R. F.) 397–452 (Academic Press, London, 1978).
2. el Shatoury, H. H. The structure of the lymph gland of *Drosophila* larvae. *Roux Arch. EntwMech. Organ.* **147,** 489–495 (1955).
3. Evans, C. J. Hartenstein, V. & Banerjee, U. Thicker than blood: conserved mechanisms in *Drosophila* and vertebrate hematopoiesis. *Dev. Cell* **5,** 673–690 (2003).
4. Jung, S. H., Evans, C. J., Uemura, C. & Banerjee, U. The *Drosophila* lymph gland as a developmental model of hematopoiesis. *Development* **132,** 2521–2533 (2005).
5. Lebestky, T., Jung, S. H. & Banerjee, U. A Serrate-expressing signaling center controls *Drosophila* hematopoiesis. *Genes Dev.* **17,** 348–353 (2003).
6. Mandal, L., Banerjee, U. & Hartenstein, V. Evidence for a fruit fly hemangioblast and similarities between lymph-gland hematopoiesis in fruit fly and mammal aorta-gonadal-mesonephros mesoderm. *Nature Genet.* **36,** 1019–1023 (2004).
7. Lo, P. C., Skeath, J. B., Gajewski, K., Schulz, R. A. & Frasch, M. Homeotic genes autonomously specify the anteroposterior subdivision of the *Drosophila* dorsal vessel into aorta and heart. *Dev. Biol.* **251,** 307–319 (2002).
8. Ryan, K. M., Hoshizaki, D. K. & Cripps, R. M. Homeotic selector genes control the patterning of seven-up expressing cells in the *Drosophila* dorsal vessel. *Mech. Dev.* **122,** 1023–1033 (2005).
9. Crozatier, M., Ubeda, J. M., Vincent, A. & Meister, M. Cellular immune response to parasitization in *Drosophila* requires the EBF orthologue collier. *PLoS Biol.* **8,** E196 (2004).
10. Casares, F. & Mann, R. S. The ground state of the ventral appendage in *Drosophila*. *Science* **293,** 1477–1480 (2001).
11. Hisa, T. *et al.* Hematopoietic, angiogenic and eye defects in Meis1 mutant animals. *EMBO J* **23,** 450–459 (2004).
12. Azcoitia, V., Aracil, M., Martinez, A. C. & Torres, M. The homeodomain protein Meis1 is essential for definitive hematopoiesis and vascular patterning in the mouse embryo. *Dev. Biol.* **280,** 307–320 (2005).
13. Schnabel, C. A., Jacobs, Y. & Cleary, M. L. HoxA9-mediated immortalization of myeloid progenitors requires functional interactions with TALE cofactors Pbx and Meis. *Oncogene* **19,** 608–616 (2000).
14. Scadden, D. T. The stem-cell niche as an entity of action. *Nature* **441,** 1075–1079 (2006).
15. Song, X., Zhu, C. H., Doan, C. & Xie, T. Germline stem cells anchored by adherens junctions in the *Drosophila* ovary niches. *Science* **296,** 1855–1857 (2002).
16. Kimble, J. & Crittenden, S. L. Germline proliferation and its control. In *WormBook* (eds The *C. elegans* Research Community), doi/10.1895/wormbook.1.13.1 (15 August, 2005).
17. Spradling, A., Drummond-Barbosa, D. & Kai, T. Stem cells find their niche. *Nature* **414,** 98–104 (2001).
18. Calvi, L. M. *et al.* Osteoblastic cells regulate the haematopoietic stem cell niche. *Nature* **425,** 841–846 (2003).
19. Zhang, J. *et al.* Identification of the haematopoietic stem cell niche and control of the niche size. *Nature* **425,** 836–841 (2003).
20. Yamashita, Y. M., Fuller, M. T. & Jones, D. L. Signaling in stem cell niches: lessons from the *Drosophila* germline. *J. Cell Sci.* **118,** 665–672 (2005).
21. Tumbar, T. *et al.* Defining the epithelial stem cell niche in skin. *Science* **303,** 359–363 (2004).
22. Abbott, M. K. & Kauffman, T. C. The relationship between the functional complexity and the molecular organization of Antennapedia locus of *Drosophila melanogaster*. *Genetics* **114,** 919–942 (1986).
23. Lum, L. & Beachy, P. A. The Hedgehog response network: sensors, switches, and routers. *Science* **304,** 1755–1759 (2004).
24. Chuang, P. T. & Kornberg, T. B. On the range of hh signaling. *Curr. Opin. Genet. Dev.* **10,** 515–522 (2002).
25. Crittenden, S. L., Leonhard, K. A., Byrd, D. T. & Kimble, J. Cellular analyses of the mitotic region in the *Caenorhabditis elegans* adult germ line. *Mol. Biol. Cell* **17,** 3051–3061 (2006).
26. Abramovich, C. & Humphries, R. K. Hox regulation of normal and leukemic hematopoietic stem cells. *Curr. Opin. Hematol.* **12,** 210–216 (2005).
27. Kieusseian, A. *et al.* Expression of Pitx2 in stromal cells is required for normal hematopoiesis. *Blood* **107,** 492–500 (2006).
28. Zhang, Y. & Kalderon, D. Hedgehog acts as a somatic stem cell factor in the *Drosophila* ovary. *Nature* **410,** 599–604 (2001).
29. Gering, M. & Patient, R. Hedgehog signaling is required for adult blood stem cell formation in zebrafish embryos. *Dev. Cell* **8,** 389–400 (2005).
30. Bhardwaj, G. *et al.* Sonic hedgehog induces the proliferation of primitive human hematopoietic cells via BMP regulation. *Nature Immunol.* **2,** 172–180 (2001).

Acknowledgements We thank R. Cripps for drawing our attention to Antennapedia expression in the developing cardiogenic mesoderm. We thank M. Crozatier, M. Meister and A. Vincent for useful discussions. We thank all members of the Banerjee and Hartenstein laboratories for their helpful comments and suggestions. In particular, we thank S. A. Sinenko for assistance with immunohistochemistry and S.-H. Jung for her initial experiments on Hedgehog. We appreciate the help of all those who provided us with reagents: W. Gehring, S. M. Cohen, L. Cooley, K. Moses, M. Crozatier, S. Noselli, A. Salzberg, H. Müller, I. Ando, K. Basler and R. Holmgren. This work was supported by an NIH grant to U.B. and V.H., and by NIH training grant support to J.A.M.-A. and C.J.E.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to U.B. (banerjee@mbi.ucla.edu).

LETTERS

Control of blood cell homeostasis in *Drosophila* larvae by the posterior signalling centre

Joanna Krzemień[1,2], Laurence Dubois[1], Rami Makki[1], Marie Meister[3]†, Alain Vincent[1] & Michèle Crozatier[1]

***Drosophila* haemocytes (blood cells) originate from a specialized haematopoietic organ—the lymph gland. Larval haematopoietic progenitors (prohaemocytes) give rise to three types of circulating haemocytes: plasmatocytes, crystal cells and lamellocytes. Lamellocytes, which are devoted to encapsulation of large foreign bodies, only differentiate in response to specific immune threats, such as parasitization by wasps. Here we show that a small cluster of signalling cells, termed the PSC (posterior signalling centre)[1], controls the balance between multipotent prohaemocytes and differentiating haemocytes, and is necessary for the massive differentiation of lamellocytes that follows parasitization. Communication between the PSC and haematopoietic progenitors strictly depends on the PSC-restricted expression of Collier, the *Drosophila* orthologue of mammalian early B-cell factor. PSC cells act, in a non-cell-autonomous manner, to maintain JAK/STAT signalling activity in prohaemocytes, preventing their premature differentiation. Serrate-mediated Notch signalling from the PSC is required to maintain normal levels of *col* transcription. The key role of the PSC in controlling blood cell homeostasis is reminiscent of interactions between haematopoietic progenitors and their microenvironment in vertebrates[2–4], thus further highlighting the interest of *Drosophila* as a model system for studying the evolution of haematopoiesis and cellular innate immunity.**

Similar to vertebrates, haematopoiesis in *Drosophila* occurs in two phases during development[5]. A first population of haemocyte precursors is specified during embryogenesis and gives rise to an invariant number of plasmatocytes and crystal cells[6]. A second population of haemocytes is specified during larval stages from a specialized haematopoietic organ, the lymph gland, which ruptures at metamorphosis and releases its prohaemocytes and haemocytes into the haemolymph[7,8]. Prohaemocytes in the lymph gland can give rise to the three types of circulating haemocytes: (1) plasmatocytes, which are involved in phagocytosis and share properties with the mammalian monocyte/macrophage lineage; (2) crystal cells, which are necessary for melanization; and (3) lamellocytes, which are devoted to encapsulation of large foreign bodies[9,10]. The lymph gland is composed of four to six paired lobes located along the aorta, with the anterior-most (primary) lobe forming first (Supplementary Fig. 2) and the posterior lobes forming later, during larval development. The primary lobes are organized in two distinct zones: a medullary zone, containing haematopoietic progenitors that can be identified by the expression of *domeless* (*dome*, which encodes the receptor for the Janus kinase/signal transducer and activator of transcription (JAK/STAT) signalling pathway), and a cortical zone, containing differentiating cells that have left the medullary zone[11]. In addition, it has been suggested that a small group of cells expressing the Notch ligand Serrate (Ser) and the transcription factor Collier (Col) acts as a 'posterior signalling centre'[1,12]. Col is required for PSC cell identity and production of lamellocytes triggered by wasp parasitism[12]. An increased number of differentiating crystal cells are found in *col* mutant lymph gland, raising the possibility that the PSC has a general role in controlling haemocyte homeostasis.

Indeed, whereas in mid-third instar wild-type larvae the medullary zone, as evidenced by *dome-gal4*/upstream activation sequence (*UAS*)-*mCD8–GFP* expression, occupies a large volume of the primary lobes (Fig. 1a), very few, if any, GFP-positive cells are detected in *col* mutant lymph gland (Fig. 1b). Expression of *tep4*, another gene specifically expressed in the medullary zone[13], is also lost in *col* mutant larvae (Fig. 1c, d), confirming a premature loss of the medullary zone. This disappearance correlates with increased differentiation both of crystal cells and plasmatocytes detected by prophenoloxidase and P1 expression, respectively (Fig. 1e, f). No difference in size or morphology of the medullary zone is observed, however, between wild type and *col* mutant larvae of second and early third instar larvae (Supplementary Fig. 3), indicating that the medullary zone forms independently of the PSC. The PSC activity is, therefore, necessary to maintain a pool of haematopoietic progenitors throughout larval development.

Whether progenitors, released from the lymph gland into the haemolymph at metamorphosis, are long-term progenitors maintaining active haematopoiesis in adult flies remains unknown owing to the lack of appropriate markers. The same phenotype as col^{-} results from selective killing of PSC cells using targeted expression of the pro-apoptotic gene *reaper* (*rpr*) (*Pcol85-Gal4/UAS-rpr*; Supplementary Fig. 1), thus strengthening the conclusion that communication between the PSC and medullary zone cells controls larval haematopoiesis. Although Notch activity within the lymph gland is required for the differentiation of crystal cells, the role of Ser in the PSC—where its expression depends on Col activity—remains uncharacterized[1,12]. To test whether Ser is involved, through Notch signalling, in the communication between the PSC and prohaemocytes, we targeted the expression of a dominant-negative form of Ser, Ser^{TM} (ref. 14), to the PSC. This resulted in a decreased level of *tep4* in the medullary zone, and increased differentiation of crystal cells (*doxA3*-positive cells), but not plasmatocytes (P1-positive cells) (data not shown), indicating that the downregulation of Ser signalling only partly reproduces the *col* mutant phenotype. Interfering with Notch signalling, either through Ser^{TM} expression in the PSC or in $Notch^{ts}$ mutants, resulted in nearly complete loss of *col* transcription in the PSC of third instar larvae, suggesting that maintaining high levels of *col* transcription in the PSC requires Notch signalling mediated by Ser (Fig. 1g, h). This observation is reminiscent of the regulatory loop involving Delta/Notch and proneural genes, which operates during lateral inhibition. Decreased *col* transcription is not due to a reduced number of PSC cells, which can be followed by GFP labelling (Fig. 1i). The number of PSC cells was normal as was their morphology, which is characterized by the extension of numerous long, thin processes (Fig. 1l) stained by

[1]Centre de Biologie du Développement, UMR 5547 and IFR 109, CNRS and Université Paul Sabatier, Toulouse III, 118 route de Narbonne, 31062 Toulouse cedex 09, France. [2]Jagiellonian University, The Institute of Zoology, 6 Ingardena Street, 30-060 Cracow, Poland. [3]UPR 9022 du CNRS, IBMC, 15 rue René Descartes, 67084 Strasbourg, France. †Present address: Museum of Zoology, 29 boulevard de la Victoire, 67000 Strasbourg, France.

an actin–GFP fusion protein (data not shown), indicating that they are filopodia. These filopodia are first observed in early third instar larvae, progressively grow in size and branching complexity throughout the third instar and could be detected to extend over 2 to 3 cell diameters (Supplementary Fig. 2). This raises the intriguing possibility that direct cellular contacts between the PSC and a subset of medullary zone cells, via filopodia, could mediate Notch signalling[15] and be important for the control of PSC function.

Col is initially expressed in all embryonic lymph gland precursor cells before becoming restricted to the PSC at the end of embryogenesis[12]. To address whether this restriction is necessary for normal haematopoiesis, we ectopically expressed *col* in medullary zone cells (*domeless-gal4/UAS-GFP/UAS-col*). Staining for markers of plasmatocytes and crystal cells revealed the nearly complete absence of haemocyte differentiation in the third instar lymph gland. Correlatively, most lymph gland cells still stained for mCD8–GFP and *tep4*, indicating that *col* expression in medullary zone cells prevents prohaemocytes from differentiating (Fig. 1j, k). In summary, experiments involving both loss of function and ectopic expression of *col* lead to the conclusion that the control of haemocyte homeostasis requires communication between two types of cells: PSC cells that segregate in late embryos and maintain high levels of *col* expression throughout larval development; and a pool of haematopoietic precursors that do not express *col*. Thus, the main role of the PSC is to block expression of genes that trigger the differentiation of prohaemocytes, a situation reminiscent of a stem cell niche[4,16]

JAK/STAT activity has been shown to promote *dome* expression in numerous tissues[17,18]. To assay whether *dome-gal4* expression in the medullary zone reflected active JAK/STAT signalling, we made use of an *in vivo* monitor for JAK/STAT activation, the *dome-MESO lacZ* transgene (*dome-MESO*)[17]. Overlap between the cells expressing high levels of *dome-GFP* and those expressing *dome-MESO* indicated that the JAK/STAT signalling pathway is activated in medullary zone cells (Fig. 2a). Similar to *dome-GFP*, expression of *dome-MESO* is independent of *col* in second instar larvae but lost in third instar larvae, confirming that *col* is required for maintaining the JAK/STAT pathway active (Supplementary Fig. 3 and data not shown). To ask whether this signalling was required to maintain prohaemocytes in the medullary zone, we analysed lymph glands from *stat92E* mutant larvae (STAT92E is the sole *Drosophila* STAT transcription factor) and found a nearly complete loss of *tep4* expression and increased haemocyte differentiation in the primary lobes (Fig. 2b–d).

These results indicate that JAK/STAT-pathway activity is required for the maintenance of the medullary zone and the control of haemocyte production in third instar larvae. The premature loss of medullary zone, together with increased differentiation of haemocytes that are observed either in the absence of a functional PSC or on loss of JAK/STAT activity, support the conclusion that PSC cells act in a non-cell-autonomous manner to maintain JAK/STAT activity in the medullary zone cells, thereby conferring on them a prohaemocyte character (Fig. 3e). We noted, however, that unlike in *col* mutant larvae, low levels of *tep4* expression could still be observed in *stat* mutant lymph glands, which correlates with morphological evidence for a residual, reduced medullary zone. Furthermore, loss of *tep4* expression and premature haemocyte differentiation in *stat92E* mutants was only observed in the primary lobes, whereas secondary lobes maintained wild-type features. In contrast, in *col* mutants, premature differentiation of crystal cells can be observed both in primary and in secondary lobes with the primary lobes remaining spherical, which is reminiscent of the morphology of lymph gland from second instar larvae. Altogether, these data suggest that the PSC performs functions in lymph gland development other than merely maintaining JAK/STAT signalling in the medullary zone. Three putative cytokine-like molecules are associated with regulating the activity of the

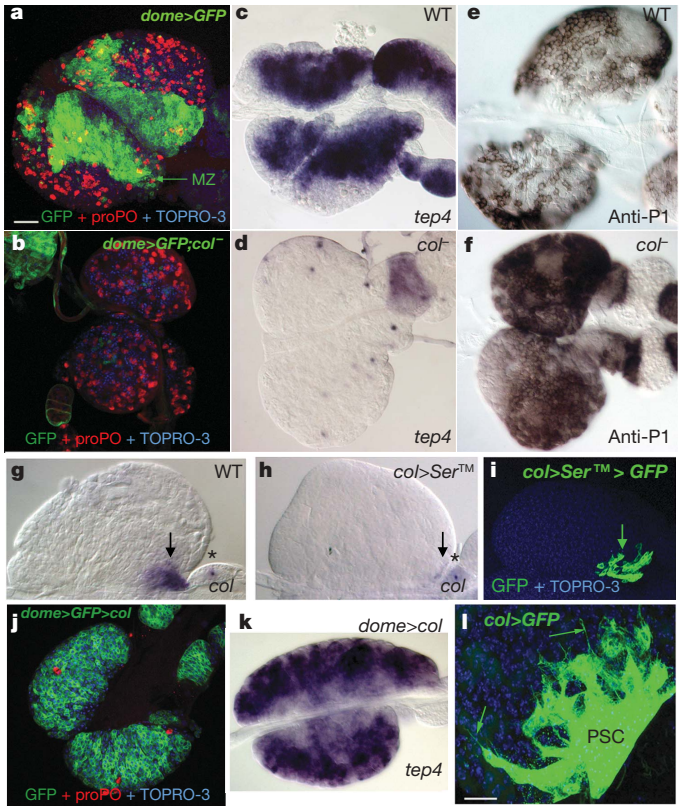


Figure 1 | The PSC is required to maintain a pool of haematopoietic progenitors. **a**–**f**, The medullary zone (marked by *dome-Gal4/UAS-mCD8–GFP* (*dome>GFP*) in **a** and **b**, and *tep4* in **c** and **d**) is lost in *col* mutant lymph glands, whereas the number of crystal cells (proPO; **a**, **b**) and plasmatocytes (P1; **e**, **f**) increases. **g**, **h**, *col* transcription is decreased in the PSC (arrow) on downregulation of Notch signalling (**h**), whereas transcription in pericardial cells (asterisk) is unmodified; the morphology of the PSC is also normal (**i**) as visualized by the expression of *mCD8–GFP* and Ser^{TM} targeted to the PSC cells (*col>*Ser^{TM}*>GFP*). **j**, **k**, Forced Col expression in the medullary zone prevents progenitors from differentiating. **l**, Expression of *mCD8–GFP* targeted to the PSC cells (*col>GFP*) reveals numerous cytoplasmic extensions (arrows). Nuclei (TOPRO-3), blue; anterior, left; scale bar, 40 μm (**a**–**k**), 20 μm (**l**).

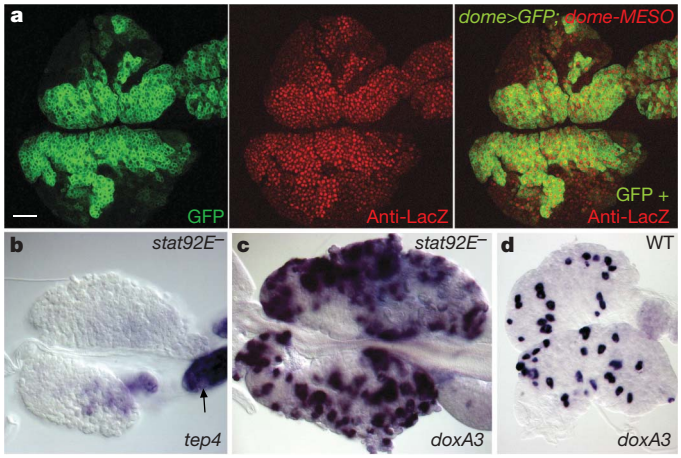


Figure 2 | JAK/STAT signalling pathway is required to maintain a pool of prohaemocytes in the lymph gland. **a**, Expression of GFP (*dome>*GFP) and nuclear LacZ (*dome-MESO*) overlaps in the medullary zone of lymph glands of third instar larvae, indicating that the JAK/STAT signalling pathway is active. Residual LacZ staining in cells that have left the medullary zone reflects the high stability of the β-galactosidase protein. **b**–**d**, In *stat* mutants (**b**, **c**), *tep4* transcription is almost completely lost from the primary lobe, correlating with increased differentiation of crystal cells. Note that in these mutants, *tep4* remains expressed in posterior lobes (arrow in **b**). Scale bar, 40 μm.

JAK/STAT pathway, Unpaired (Upd) 1 to 3. Overexpression of either *upd1* or *upd2* in the PSC (using the *Pcol85-Gal4* driver) or of an active form of the JAK kinase ($hop^{Tum\text{-}l}$) in the medullary zone (using a *dome-gal4* driver) provokes larval lethality at the L1/L2 stage, precluding analysis at the L3 stage (data not shown). Which Upd(s) acts in the medullary zone under the control of the PSC remains an open question.

Egg-laying by parasitoid wasps in *Drosophila* larvae triggers the production of lamellocytes. Using antibodies raised against the lamellocyte-specific integrin α-chain, α-PS4 (refs 12, 13), we confirmed the absence of lamellocytes in unchallenged larvae. In contrast, massive differentiation of lamellocytes was already observed 24 h after wasp egg-laying (Fig. 3c, d). Strikingly, this was paralleled by the premature loss of the medullary zone (Fig. 3a, b), indicating that the production of lamellocytes takes place at the expense of the pool of prohaemocytes, which, during normal development, are still present in the lymph gland when it bursts at metamorphosis. This reprogramming, bypassing the normal haemocyte homeostasis, reveals an unexpected level of plasticity of *Drosophila* prohaemocytes, which, to our knowledge, are the only *Drosophila* cells with a 'cryptic' fate revealed under only specific conditions. Because lamellocyte differentiation could not be observed when prohaemocytes differentiated prematurely, for instance, in the absence of the PSC (see Fig. 1) or in *hopscotch* mutants (*hop*, the *Drosophila* JAK) after wasp infestation (data not shown), it indicates that the maintenance of a pool of multipotent progenitors in the lymph gland is a prerequisite for *Drosophila* to be able to mount a dedicated cellular response against parasites (Fig. 3f). Conversely, the premature loss of haemocyte precursors that results from wasp infestation could have a cost in terms either of adult fitness and/or reproductive ability[19,20].

The key role of the PSC in maintaining a pool of uncommitted haematopoietic progenitors throughout *Drosophila* larval development that is required to generate the necessary supply of lamellocytes in response to parasitization is very reminiscent of the micro-environmental niches that supply factors to maintain stem cell potential in vertebrates[2,3] and in the *Drosophila* germ line[21]. A signal produced by the PSC is required to maintain JAK/STAT activity and haemocyte progenitors in medullary zone cells (Fig. 3e). This is also reminiscent of the key role of JAK/STAT signalling in the stem cell niches of *Drosophila* male and female gonads[21,22]. Expression of Ser in the PSC, and the role of Notch signalling in the control of germline stem cell proliferation in nematodes and haematopoietic stem cell function in vertebrates, suggest some further parallels[23–25]. Yet, evidence for the existence of bona fide haematopoietic stem cells in the *Drosophila* lymph gland is still lacking. Further characterization of PSC functions and related signalling components in the control of *Drosophila* larval haematopoiesis should provide important new insights into the interactions between the haematopoietic progenitors and their micro-environment, and the evolution of cellular aspects of innate immunity.

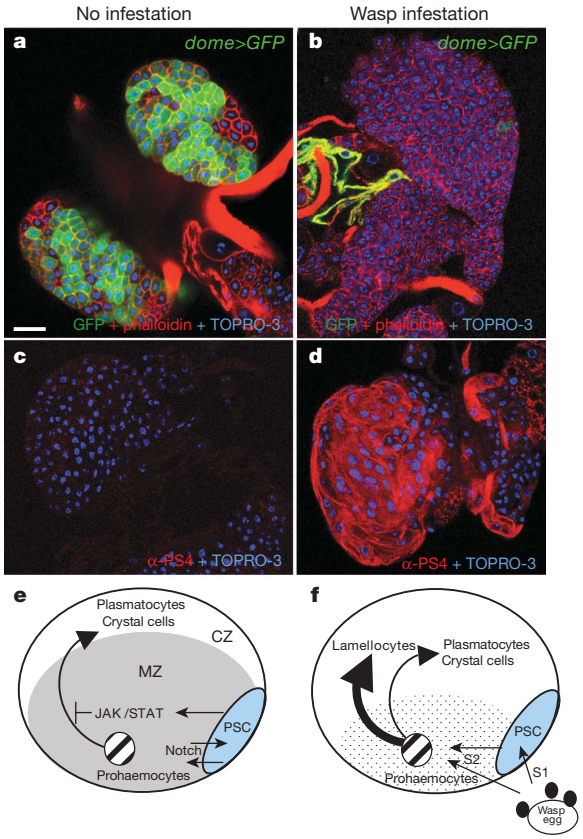


Figure 3 | Premature differentiation of prohaemocytes into lamellocytes on wasp infestation. **a**, **b**, The medullary zone is lost after wasp infestation and this is paralleled by lamellocyte differentiation (α-PS4) (**c**, **d**). Cell contours are revealed by phalloidin staining (red in panels **a** and **b**). Scale bar, 20 μm. **e**, **f**, Model for larval haematopoiesis. **e**, During normal development the PSC acts non-cell-autonomously to maintain JAK/STAT signalling activity in the medullary zone (grey), thereby preventing premature haemocyte differentiation. Notch signalling maintains *col* expression in the PSC. **f**, On parasitization, a signal, probably emitted by plasmatocytes (black circles), is perceived either directly (S2) or via the PSC (S1→S2 relay), or through both ways. Col activity in the PSC is required for pro-haemocytes to adopt a lamellocyte fate. Arrows, activation; vertical bars, repression.

METHODS

Genetics and *Drosophila* strains. A lethal *P-lacZ* insertion (*Pcol85-lacZ*) located immediately upstream of the *col* transcription start site (position −624 base pairs from +1 of transcription) from which LacZ expression mimics that of *col* was generously provided by W. Son. To convert it into a driver line, we used the technique of gene replacement[26] and substituted the *Gal4* for the *lacZ* coding region, thereby generating a *Gal4* line, *Pcol85-Gal4*. In *Pcol85-Gal4/UAS-lacZ* strains LacZ expression is first detected in the lymph gland at embryonic stage 16 and is restricted to PSC cells from that stage until pupariation (Supplementary Fig. 1). The $Oregon^{R}$ strain was used as wild type; *Pdome-Gal4* was described in ref. 27. Wasp infestation was as in ref. 12. To kill the PSC cells, *Pcol85-Gal4* was used to target the expression of *rpr* in the PSC; embryos were maintained at 16 °C to avoid embryonic lethality and larval development was at 29 °C. We used a null allele of *stat*, $stat^{397}$. Larval development of *Pcol85-Gal4/UAS-*Ser^{TM} was performed at 25 °C. Embryonic and larval development of $Notch^{ts2}$ mutant occurred at 29 °C.

Antibodies. Immunostaining and *in situ* hybridization were performed as in ref. 28. To produce polyclonal antibodies against α-PS4 we used a synthetic peptide of 13 amino acids (Leu-Lys-Leu-Ala-Asp-Thr-Asn-Leu-Iso-Pro-Tyr-Tyr-Ser) coupled to ovalbumin (Imject Maleimide activated Ovalbumin, Pierce) to immunize rabbits. A dilution of 1/200 was used for immunostaining on lymph glands. The following antibodies, dilutions and reagents were used: anti-mouse and anti-rabbit anti-lacZ (Promega, 1/1,000, and Capell, 1/5,000, respectively); rabbit poly-clonal anti-prophenoloxydase (1/200) and mouse mono-clonal anti-P1 (1/30) (gifts from H.M. Müeller and I. Ando, respectively), rabbit anti-Col[28], and rabbit poly-clonal anti-GFP (Torrey Pines Biolabs, 1/1000). Alexa-Fluor-546-conjugated goat anti-mouse (Molecular Probes, 1/250), Alexa-Fluor-647-conjugated goat anti-mouse (Molecular Probes, 1/250) biotinylated anti-mouse and anti-rabbit (Vector, 1/800), phalloidin (Interchim) and TOPRO3 (Molecular Probes).

Received 24 August 2006; accepted 5 February 2007.

1. Lebestky, T., Jung, S. H. & Banerjee, U. A Serrate-expressing signaling center controls *Drosophila* hematopoiesis. *Genes Dev.* **17**, 348–353 (2003).
2. Nagasawa, T. Microenvironmental niches in the bone marrow required for B-cell development. *Nature Rev. Immunol.* **6**, 107–116 (2006).
3. Wilson, A. & Trumpp, A. Bone-marrow haematopoietic-stem-cell niches. *Nature Rev. Immunol.* **6**, 93–106 (2006).
4. Scadden, D. T. The stem-cell niche as an entity of action. *Nature* **441**, 1075–1079 (2006).
5. Holz, A., Bossinger, B., Strasser, T., Janning, W. & Klapper, R. The two origins of haemocytes in *Drosophila*. *Development* **130**, 4955–4962 (2003).

6. Bataille, L., Auge, B., Ferjoux, G., Haenlin, M. & Waltzer, L. Resolving embryonic blood cell fate choice in *Drosophila*: interplay of GCM and RUNX factors. *Development* **132,** 4635–4644 (2005).
7. Rizki, R. M. & Rizki, T. M. Selective destruction of a host blood cell type by a parasitoid wasp. *Proc. Natl Acad. Sci. USA* **81,** 6154–6158 (1984).
8. Lanot, R., Zachary, D., Holder, F. & Meister, M. Postembryonic hematopoiesis in *Drosophila. Dev. Biol.* **230,** 243–257 (2001).
9. Evans, C. J., Hartenstein, V. & Banerjee, U. Thicker than blood: conserved mechanisms in *Drosophila* and vertebrate hematopoiesis. *Dev. Cell* **5,** 673–690 (2003).
10. Meister, M. & Lagueux, M. *Drosophila* blood cells. *Cell. Microbiol.* **5,** 573–580 (2003).
11. Jung, S. H., Evans, C. J., Uemura, C. & Banerjee, U. The *Drosophila* lymph gland as a developmental model of hematopoiesis. *Development* **132,** 2521–2533 (2005).
12. Crozatier, M., Ubeda, J. M., Vincent, A. & Meister, M. Cellular immune response to parasitization in *Drosophila* requires the EBF orthologue Collier. *PLoS Biol.* **2,** E196 (2004).
13. Irving, P. *et al.* New insights into *Drosophila* larval haemocyte functions through genome-wide analysis. *Cell. Microbiol.* **7,** 335–350 (2005).
14. Sun, X. & Artavanis-Tsakonas, S. The intracellular deletions of Delta and Serrate define dominant negative forms of the *Drosophila* Notch ligands. *Development* **122,** 2465–2474 (1996).
15. De Joussineau, C. *et al.* Delta-promoted filopodia mediate long-range lateral inhibition in *Drosophila. Nature* **426,** 555–559 (2003).
16. Spradling, A., Drummond-Barbosa, D. & Kai, T. Stem cells find their niche. *Nature* **414,** 98–104 (2001).
17. Hombria, J. C., Brown, S., Hader, S. & Zeidler, M. P. Characterisation of Upd2, a *Drosophila* JAK/STAT pathway ligand. *Dev. Biol.* **288,** 420–433 (2005).
18. Arbouzova, N. I. & Zeidler, M. P. JAK/STAT signalling in *Drosophila*: insights into conserved regulatory and cellular functions. *Development* **133,** 2605–2616 (2006).
19. Kraaijeveld, A. R. & Godfray, H. C. Trade-off between parasitoid resistance and larval competitive ability in *Drosophila melanogaster. Nature* **389,** 278–280 (1997).
20. Rolff, J. & Siva-Jothy, M. T. Invertebrate ecological immunology. *Science* **301,** 472–475 (2003).
21. Decotto, E. & Spradling, A. C. The *Drosophila* ovarian and testis stem cell niches: similar somatic stem cells and signals. *Dev. Cell* **9,** 501–510 (2005).
22. Lin, H. The stem-cell niche theory: lessons from flies. *Nature Rev. Genet.* **3,** 931–940 (2002).
23. Kimble, J. & Crittenden, S. L. Germline proliferation and its control. In *WormBook* (eds The *C. elegans* Research Community), doi/10.1895/worm-book.1.13.1 (15 August, 2005).
24. Calvi, L. M. *et al.* Osteoblastic cells regulate the haematopoietic stem cell niche. *Nature* **425,** 841–846 (2003).
25. Zhang, J. *et al.* Identification of the haematopoietic stem cell niche and control of the niche size. *Nature* **425,** 836–841 (2003).
26. Sepp, K. J. & Auld, V. J. Conversion of *lacZ* enhancer trap lines to *GAL4* lines using targeted transposition in *Drosophila melanogaster. Genetics* **151,** 1093–1101 (1999).
27. Bourbon, H. M. *et al.* A P-insertion screen identifying novel X-linked essential genes in *Drosophila. Mech. Dev.* **110,** 71–83 (2002).
28. Crozatier, M. & Vincent, A. Requirement for the *Drosophila* COE transcription factor Collier in formation of an embryonic muscle: transcriptional response to Notch signalling. *Development* **126,** 1495–1504 (1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Castelli-Gair Hombria, O. Devergne, M. Milan, S. Noselli, M. Zeidler and the Bloomington Stock Center for fly stocks, I. Ando and H. M. Müller for antibodies, and M. Haenlin, J. Smith, L. Waltzer for critical reading of the manuscript and discussion. We are grateful to platforme Toulouse RIO imaging and B. Ronsin for assistance with confocal microscopy. This work was supported by CNRS, a European Marie Curie PhD Training programme grant and Ministère de la Recherche et de la technologie (ACI Biologie Cellulaire, Moléculaire et Structurale) et Ministère des affaires étrangères (Programme Egide).

Author Contributions J.K., M.M., A.V. and M.C. conceived the experiments, which were essentially performed by J.K. and M.C. L.D. and M.M. provided crucial reagents. R.M. assisted with experiments involving wasp infestation. A.V. and M.C. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.C. (crozat@cict.fr).

LETTERS

14-3-3σ controls mitotic translation to facilitate cytokinesis

Erik W. Wilker[1], Marcel A. T. M. van Vugt[1], Stephen C. Artim[1], Paul H. Huang[2], Christian P. Petersen[1], H. Christian Reinhardt[1], Yun Feng[1], Phillip A. Sharp[1], Nahum Sonenberg[3], Forest M. White[2] & Michael B. Yaffe[1,2]

14-3-3 proteins are crucial in a wide variety of cellular responses including cell cycle progression, DNA damage checkpoints and apoptosis. One particular 14-3-3 isoform, σ, is a p53-responsive gene, the function of which is frequently lost in human tumours, including breast and prostate cancers as a result of either hypermethylation of the 14-3-3σ promoter or induction of an oestrogen-responsive ubiquitin ligase that specifically targets 14-3-3σ for proteasomal degradation[1–9]. Loss of 14-3-3σ protein occurs not only within the tumours themselves but also in the surrounding pre-dysplastic tissue (so-called field cancerization), indicating that 14-3-3σ might have an important tumour suppressor function that becomes lost early in the process of tumour evolution[3,9]. The molecular basis for the tumour suppressor function of 14-3-3σ is unknown. Here we report a previously unknown function for 14-3-3σ as a regulator of mitotic translation through its direct mitosis-specific binding to a variety of translation/initiation factors, including eukaryotic initiation factor 4B in a stoichiometric manner. Cells lacking 14-3-3σ, in marked contrast to normal cells, cannot suppress cap-dependent translation and do not stimulate cap-independent translation during and immediately after mitosis. This defective switch in the mechanism of translation results in reduced mitotic-specific expression of the endogenous internal ribosomal entry site (IRES)-dependent form of the cyclin-dependent kinase Cdk11 (p58 PITSLRE), leading to impaired cytokinesis, loss of Polo-like kinase-1 at the midbody, and the accumulation of binucleate cells. The aberrant mitotic phenotype of 14-3-3σ-depleted cells can be rescued by forced expression of p58 PITSLRE or by extinguishing cap-dependent translation and increasing cap-independent translation during mitosis by using rapamycin. Our findings show how aberrant mitotic translation in the absence of 14-3-3σ impairs mitotic exit to generate binucleate cells and provides a potential explanation of how 14-3-3σ-deficient cells may progress on the path to aneuploidy and tumorigenesis.

Because many tumour suppressor proteins function during specific parts of the cell cycle, we examined whether the binding of ligands to 14-3-3σ showed a cell cycle dependence. U2OS cells were synchronized by a double thymidine block, and lysates were prepared at various times after release (Supplementary Fig. 1a). Immunoprecipitation of endogenous 14-3-3σ from these lysates showed a marked increase in 14-3-3σ-bound proteins in the mitotic and immediate post-mitotic periods (Fig. 1a, time points at 12 and 16 h after release). This same mitotic enrichment in ligand binding to 14-3-3σ was also observed in other cell types, including HCT116 cells and HeLa cells (Supplementary Fig. 1b–d), but was not observed when ligands of another endogenous 14-3-3 isoform, 14-3-3β, were analysed (Supplementary Fig. 1e) or in 14-3-3σ immunoprecipitations from asynchronous cells after DNA damage (Supplementary Fig. 2). The mitosis-dependent binding of ligands to 14-3-3σ seemed to be both dependent on phosphorylation, because binding was lost when lysates were treated with protein phosphatase 1 (Supplementary Fig. 1d), and direct, as shown with a two-dimensional 'far-western' assay (Fig. 1d).

The mitotic targets of 14-3-3σ were identified by mass spectrometry. This analysis revealed that a large proportion of 14-3-3σ ligands were proteins involved in the process of translation, including several initiation factors involved in mediating cap-dependent translation (Fig. 1b and Supplementary Table 1). Two-dimensional far-western blotting of mitotic 14-3-3σ immunoprecipitates showed that several of these proteins, including eukaryotic initiation factor 4B (eIF4B), eIF-2α and elongation factor 1α (EF1α), bound directly to 14-3-3σ, whereas other 14-3-3σ-associated proteins such as eIF4G seemed to be interacting indirectly (Fig. 1e).

It is well established that translation in mammalian cells is markedly altered during and immediately after mitosis[10], with a pronounced suppression of cap-dependent translation and a corresponding enhancement of cap-independent translation[11]. The biological mechanism and significance of this translational switch, however, are unclear. Given the abundance of proteins involved in translation in the 14-3-3σ immunoprecipitations, we investigated whether 14-3-3σ has a direct function in this process by using RNA interference (RNAi). Synchronous 14-3-3σ stable knockdown and control cells were pulse-labelled with [^{35}S]methionine, and newly translated proteins were analysed by SDS–polyacrylamide gel electrophoresis (PAGE) and autoradiography (Fig. 2a, d, and Supplementary Fig. 3a). As expected, in control short hairpin RNA (shRNA)-treated HeLa and U2OS cells, radiolabel incorporation into nascent polypeptides was suppressed at 12 and 16 h after release from the double thymidine block (Fig. 2a, e, lanes marked 'con')[11]. Remarkably, no suppression of mitotic translation was observed in either of these cell types when they were depleted of 14-3-3σ (lanes marked 'σ'). Identical results were also obtained with two other distinct 14-3-3σ shRNAs (data not shown). In contrast to 14-3-3σ depletion, RNAi-mediated knockdown of 14-3-3β had no effect on the suppression of translation during and immediately after mitosis (Fig. 2b, d). Furthermore, the aberrant mitotic translation seen in the 14-3-3σ knockdown cells did not result from a failure of these cells to enter mitosis (Fig. 2c). Although the global pattern of cell cycle progression did not change, we observed an increase in the population of 4*n* DNA-containing cells and a significant increase in mitotic index (Fig. 2c, and Supplementary Fig. 3b, c). Coexpression of an shRNA-resistant construct of 14-3-3σ together with 14-3-3σ shRNA restored the suppression of mitotic translation (Fig. 2e, lanes marked 'σ_r', and Fig. 2f, top). These findings indicate that 14-3-3σ is important

[1]Center for Cancer Research, Department of Biology and [2]Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. [3]Department of Biochemistry, McGill Cancer Centre, McGill University, Montreal, Quebec, H3G 1Y6, Canada.

in the physiological downregulation of new protein synthesis during and immediately after mitosis.

The inhibition of translation observed during mitosis has previously been shown to result from a global decrease in cap-dependent translation, whereas cap-independent translation is subsequently increased[12,13]. We used two different IRES-containing bi-cistronic vectors, a viral-based IRES from HIV[14], and a normal cellular IRES from $p27^{Kip1}$ (ref. 15), in a direct investigation of the relative effects of 14-3-3σ on cap-dependent versus cap-independent translation in synchronized U2OS and HeLa cells. As shown in Fig. 2g, 14-3-3σ-depleted cells failed to show the robust mitotic-specific increase in the ratio of cap-independent to cap-dependent translation that occurred in control cells. Thus, a loss of 14-3-3σ resulted in both aberrant persistence of cap-dependent translation and prevented the normal enhancement of cap-independent to cap-dependent translation during mitosis.

The relative amounts of different eukaryotic initiation factors in cells varies more than 100-fold[16], and they have distinct roles in the cap recognition process[17,18]. To investigate which eIF targets of 14-3-3σ are potentially responsible for the 14-3-3σ-mediated suppression of cap-dependent mitotic translation, we immunodepleted mitotic extracts with the use of anti-14-3-3σ antibodies and probed the depleted lysates for residual eIFs. Of the eIFs examined, the most pronounced result was obtained with eIF4B, which was completely absent from the 14-3-3σ-depleted mitotic lysates (Fig. 1c). eIF4B facilitates the ATP-dependent helicase activity of eIF4A to promote ribosome recruitment required for cap-dependent translation[19]. We further observed that transient overexpression of exogenous Flag-tagged eIF4B was able to overcome the ability of endogenous 14-3-3σ to suppress mitotic translation in normal cells but had no effect on the increased translation seen in 14-3-3σ knockdown cells (Fig. 2e, lanes marked '4B', and Fig. 2f, bottom). Taken together, the data in Figs 1 and 2 indicate that 14-3-3σ is important in the physiological downregulation of new protein synthesis during and immediately after mitosis by suppressing cap-dependent translation through binding to eIF4B.

Because loss of 14-3-3σ occurs early in the process of tumorigenesis, we also examined the 14-3-3σ knockdown cells for a corresponding cellular phenotype that might correlate with aberrant regulation of mitotic translation. In the 14-3-3σ knockdown cells we observed significantly increased numbers of cells displaying persistent cytokinetic bridges, binucleate cells (often containing midbody remnants), and cells that seemed to have 'fused' after mitosis with widely separated nuclei, probably reflecting a failure late in cytokinesis (Fig. 3a–c; $P < 0.05$ for 14-3-3σ shRNA versus control cells, Student's *t*-test, two-tailed; and Supplementary Fig. 3d–f). Intriguingly, these phenotypes seemed to correlate with a failure of Polo-like kinase-1 (Plk1), a critical mitotic kinase involved in the completion of cytokinesis[20,21], to localize to the midbody at the end of mitosis (Fig. 4a). Both the localization of Plk1 to the midbody and the mitotic phenotype of 14-3-3σ-depleted cells could be significantly restored by treatment of the cells with rapamycin, an inhibitor of cap-dependent, but not cap-independent, translation[22], immediately before mitotic entry (Figs 3b, c and 4a).

We searched for proteins that were translated in mitosis in a cap-independent manner and were postulated to be involved in the mitotic process. One of the candidates, PITSLRE/Cdk11, is a member of the Cdc2-like protein kinase family that undergoes cap-independent translation from an internal ribosome entry site during mitosis to produce a 58-kDa isoform that facilitates proper mitotic progression and termination[23–25]. In both U2OS cells and HeLa cells we observed that 14-3-3σ knockdown cells, but not control cells, failed to synthesize p58 PITSLRE in mitosis (Fig. 4b, c, 16 h lanes). Depletion of PITSLRE/Cdk11 from U2OS cells caused an increase in mitotic content as well as the number of binucleate cells, exactly as observed in 14-3-3σ-depleted cells (Supplementary Figs 3 and 4). Addition of rapamycin to 14-3-3σ knockdown cells restored the mitotic translation of p58 PITSLRE (Fig. 4b, c). Furthermore, transient transfection of a complementary DNA encoding the 58-kDa isoform of PITSLRE into the 14-3-3σ knockdown cells was sufficient to relocalize Plk1 at the midbody and partly rescue the mitosis-defective cell phenotype, reducing the number of binucleate and fused cells by slightly more

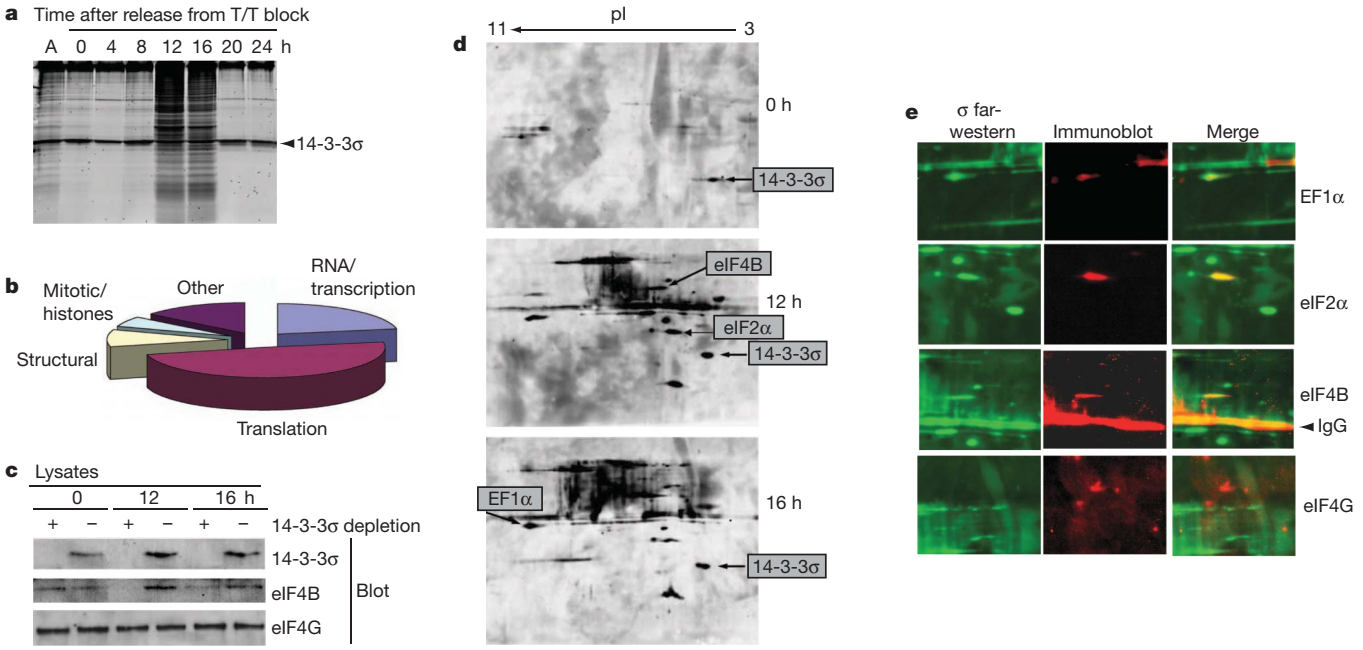


Figure 1 | 14-3-3σ binds to its targets during mitosis. **a**, SDS–PAGE analysis of 14-3-3σ immunoprecipitates from synchronized U2OS cells. Sypro-Ruby stain. T/T, double thymidine; A, asynchronous cells. **b**, Mitotic ligands of 14-3-3σ identified by mass spectrometry categorized into five major groups. **c**, 14-3-3σ quantitatively immunodepletes eIF4B but not eIF4G from mitotic (12 h), but not from interphase (0 h), U2OS cell extracts. **d**, Two-dimensional far-western blotting of 14-3-3σ immunoprecipitates, with purified 14-3-3σ as a probe for direct binding. pI, isoelectric point. **e**, eIF4B, eIF2α and EF1α as direct mitotic ligands of 14-3-3σ. Two-dimensional far-western blots (green) re-probed with antibodies against the indicated proteins (red). eIF4G in the 14-3-3σ immunoprecipitates does not directly bind to 14-3-3σ.

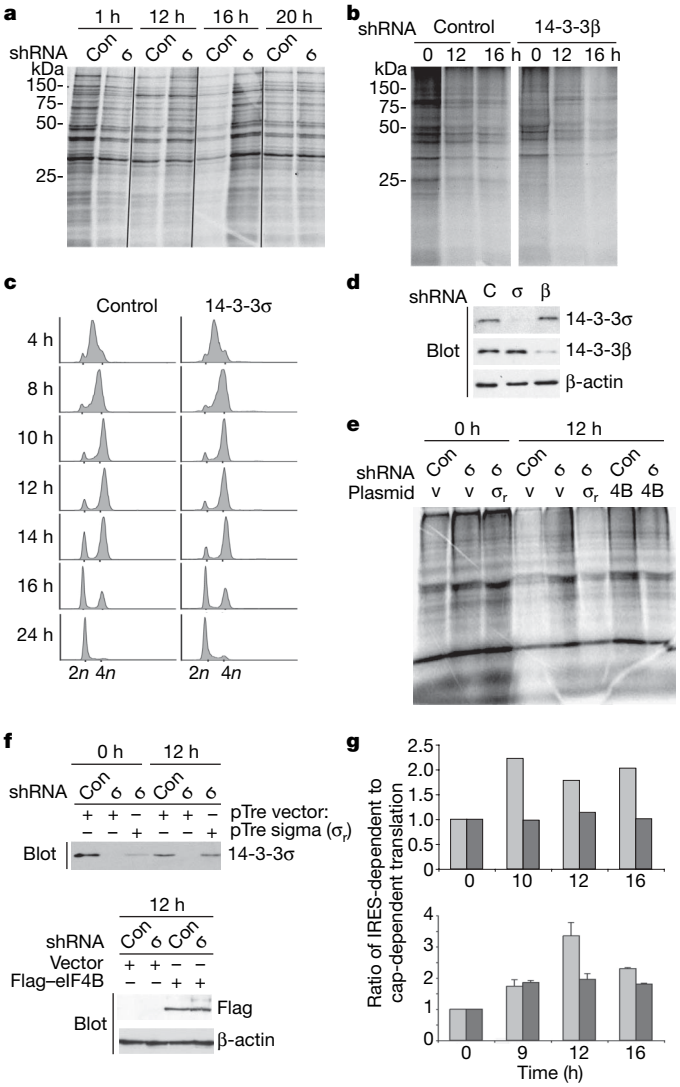


Figure 2 | 14-3-3σ knockdown cells fail to suppress cap-dependent translation during and immediately after mitosis. **a**, **b**, Control, 14-3-3σ (**a**) and 14-3-3β (**b**) shRNA HeLa cells were pulse-labelled with [^{35}S]methionine. New protein synthesis was assayed by SDS–PAGE and autoradiography. Numbers at the left indicate molecular masses. **c**, Cell cycle distribution of control and 14-3-3σ shRNA-treated cells. **d**, Efficiency of 14-3-3 knockdown. C, control. **e**, Overexpression of eIF4B overcomes 14-3-3σ-dependent suppression of translation in mitosis. V, vector; 4B, Flag-eIF4B; σ_r, RNAi-resistant 14-3-3σ. Synchronized cells pulse-labelled as in **b** at indicated times. **f**, Expression of σ_r and Flag–eIF4B in **e**. **g**, Ratio of cap-independent to cap-dependent translation in synchronized control (light grey bars) and 14-3-3σ (dark grey bars) knockdown cells assayed by firefly/*Renilla* luciferase activity at the indicated times. Top, U2OS cells and HIV IRES; bottom, HeLa cells and p27^{Kip1} IRES. Results are means and s.e.m. from duplicate experiments.

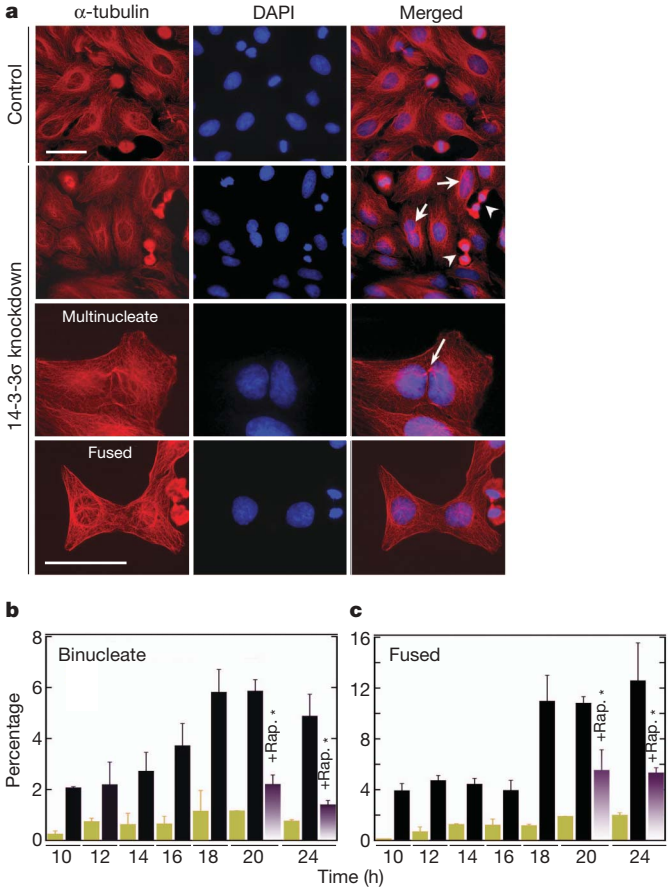


Figure 3 | Depletion of 14-3-3σ results in impaired cytokinesis. **a**, 14-3-3σ knockdown U2OS cells, stained as indicated, show increased numbers of binucleate cells (short arrows) and cells in terminal cytokinesis (arrowheads). Close-up views demonstrate persistent midbody structures (long arrow), and a 'fused' post-mitotic cell lacking distinct intercellular boundaries. Scale bars, 10 μm. DAPI, 4,6-diamidino-2-phenylindole. **b**, **c**, Rapamycin treatment before mitotic entry suppresses binucleate (**b**) and fused-cell (**c**) formation in 14-3-3σ knockdown cells. Synchronized U2OS shRNA cells were scored at the indicated times after release. Green bars, control shRNA; black bars, 14-3-3σ shRNA; +Rap., 14-3-3σ shRNA cells treated with 20 ng ml^{-1} rapamycin at 9 h. Asterisk, $P < 0.02$ for rapamycin-treated versus untreated 14-3-3σ knockdown cells. Results are means and s.d. from three independent experiments.

than 50% (Fig. 4d, e, and Supplementary Fig. 4i). Thus, 14-3-3σ is required in normal mitosis to suppress cap-dependent translation through binding eIF4B, allowing the cap-independent translation of critical mitotic regulators including p58 PITSLRE kinase.

Two recent reports[26,27] have suggested that the formation of binucleate cells as a result of cytokinesis failure is an early event in tumour formation and underlies the subsequent development of genomic instability. We have observed that 14-3-3σ, an important tumour suppressor protein whose expression is lost in a variety of epithelial tumours, has a critical function in regulating protein synthesis during and immediately after mitosis and is required for the IRES-dependent translation of p58 PITSLRE, a protein kinase critical to the proper completion of cytokinesis. These findings therefore establish a connection between the aberrant regulation of mitotic translation and improper cytokinesis resulting in a phenotype that is associated with early stages of human oncogenesis.

METHODS

Immunoprecipitation studies and mass spectrometry analysis. U2OS, HeLa and HCT116 cell lysates were prepared and immunoprecipitated overnight at 4 °C with an anti-14-3-3σ monoclonal antibody (CS112)[28,29]. Immunoprecipitates were digested with trypsin and analysed on a QSTAR XL Pro quadrupole time-of-flight mass spectrometer (Applied Biosystems).

Two-dimensional gel electrophoresis. 14-3-3σ immunoprecipitates were resolved by isoelectric focusing with IPGphor strips as described previously[29]. Separation in the second dimension was performed with 12% Anderson gels and used for far-western blot analyses with bacterially purified 14-3-3σ (about 50 μg ml^{-1}) as a probe, followed by immunoblotting with CS112.

Lentiviral shRNA constructs. 14-3-3σ, 14-3-3β and p58 PITSLRE shRNAs were subcloned into pLentilox 3.7 for lentiviral production as described[30]. Heterogeneous populations stably expressing shRNA were established in HeLa and U2OS cells.

[^{35}S]Methionine pulse-labelling. Cells were labelled with 100 μCi of [^{35}S]methionine for 20 min, lysed, and precipitated with acetone. Resuspended

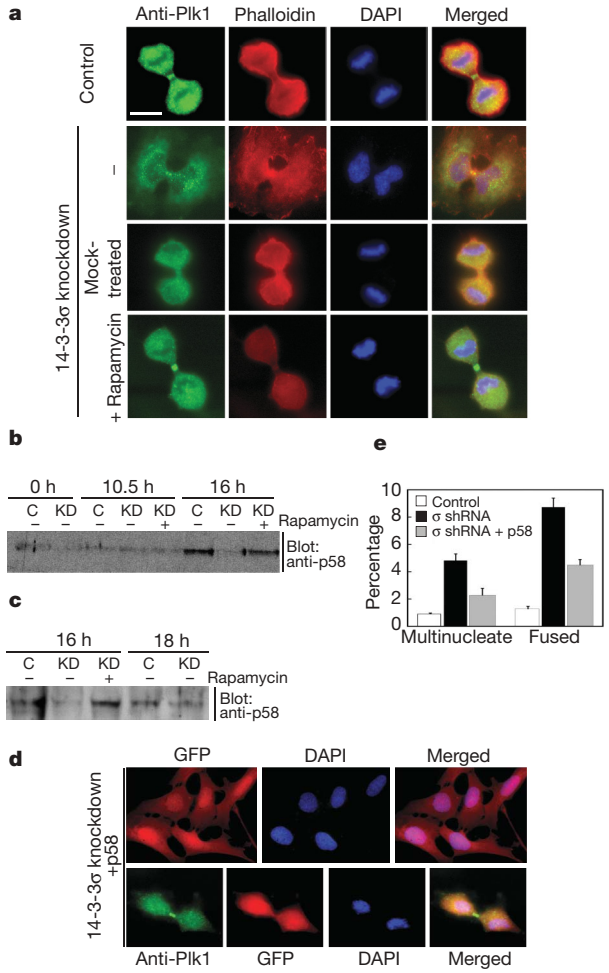


Figure 4 | Depletion of 14-3-3σ blocks IRES-dependent mitotic translation of p58 PITSLRE kinase and results in failure of Plk1 to localize at the midbody. **a**, Lack of localization of Plk1 to the midbody and cell fusion with failure to complete cytokinesis in 14-3-3σ knockdown U2OS cells is reversed by rapamycin before mitotic entry. Scale bar, 5 μm. **b**, **c**, Synchronized control (C) or 14-3-3σ shRNA-treated (KD) U2OS (**b**) and HeLa (**c**) cells, with and without rapamycin treatment, lysed at the indicated times and immunoblotted for endogenous p58 PITSLRE expression[25]. IRES-dependent translation of p58 PITSLRE in 14-3-3σ knockdown cells was restored by rapamycin. **d**, **e**, Forced expression of p58 PITSLRE reverts the mitotic phenotype of 14-3-3σ knockdown cells. GFP-expressing 14-3-3σ knockdown U2OS cells were transfected with p58 PITSLRE or vector control, then imaged (**d**) and scored (**e**) 18 h after release from a double-thymidine block. Results in **e** are means and s.d. from three independent experiments.

proteins were separated on 12% SDS–PAGE gels and analysed by autoradiography. Where indicated, cells were treated with rapamycin at a final concentration of 20 ng ml^{-1} 9 h after release from a double thymidine block.

Additional information and detailed protocols are provided in Supplementary Methods.

Received 13 July 2006; accepted 9 January 2007; corrected 10 April 2007.

1. Vercoutter-Edouart, A. S. *et al.* Proteomic analysis reveals that 14-3-3σ is down-regulated in human breast cancer cells. *Cancer Res.* **61**, 76–80 (2001).
2. Ferguson, A. T. *et al.* High frequency of hypermethylation at the 14-3-3σ locus leads to gene silencing in breast cancer. *Proc. Natl Acad. Sci. USA* **97**, 6049–6054 (2000).
3. Umbricht, C. B. *et al.* Hypermethylation of 14-3-3σ (stratifin) is an early event in breast cancer. *Oncogene* **20**, 3348–3353 (2001).
4. Urano, T. *et al.* Efp targets 14-3-3σ for proteolysis and promotes breast tumour growth. *Nature* **417**, 871–875 (2002).
5. Moreira, J. M., Gromov, P. & Celis, J. E. Expression of the tumor suppressor protein 14-3-3σ is down-regulated in invasive transitional cell carcinomas of the urinary bladder undergoing epithelial-to-mesenchymal transition. *Mol. Cell. Proteomics* **3**, 410–419 (2004).
6. Iwata, N. *et al.* Frequent hypermethylation of CpG islands and loss of expression of the 14-3-3σ gene in human hepatocellular carcinoma. *Oncogene* **19**, 5298–5302 (2000).
7. Suzuki, H. *et al.* Inactivation of the 14-3-3σ gene is associated with 5′ CpG island hypermethylation in human cancers. *Cancer Res.* **60**, 4353–4357 (2000).
8. Villaret, D. B. *et al.* Identification of genes overexpressed in head and neck squamous cell carcinoma using a combination of complementary DNA subtraction and microarray analysis. *Laryngoscope* **110**, 374–381 (2000).
9. Gasco, M. *et al.* Coincident inactivation of 14-3-3σ and p16INK4a is an early event in vulval squamous neoplasia. *Oncogene* **21**, 1876–1881 (2002).
10. Prescott, A. R. Synthesis of RNA and protein during mitosis in mammalian tissue culture cells. *Exp. Cell Res.* **26**, 260–268 (1962).
11. Pyronnet, S., Dostie, J. & Sonenberg, N. Suppression of cap-dependent translation in mitosis. *Genes Dev.* **15**, 2083–2093 (2001).
12. Pyronnet, S., Pradayrol, L. & Sonenberg, N. A cell cycle-dependent internal ribosome entry site. *Mol. Cell* **5**, 607–616 (2000).
13. Qin, X. & Sarnow, P. Preferential translation of internal ribosome entry site-containing mRNAs during the mitotic cycle in mammalian cells. *J. Biol. Chem.* **279**, 13721–13728 (2004).
14. Brasey, A. *et al.* The leader of human immunodeficiency virus type 1 genomic RNA harbors an internal ribosome entry segment that is active during the G2/M phase of the cell cycle. *J. Virol.* **77**, 3939–3949 (2003).
15. Kullmann, M., Gopfert, U., Siewe, B. & Hengst, L. B. ELAV/Hu proteins inhibit p27 translation via an IRES element in the p27 5′UTR. *Genes Dev.* **16**, 3087–3099 (2002).
16. Hershey, J. W. Expression of initiation factor genes in mammalian cells. *Biochimie* **76**, 847–852 (1994).
17. Holcik, M. & Sonenberg, N. Translational control in stress and apoptosis. *Nature Rev. Mol. Cell Biol.* **6**, 318–327 (2005).
18. Merrick, W. C. Cap-dependent and cap-independent translation in eukaryotic systems. *Gene* **332**, 1–11 (2004).
19. Rogers, G. W. Jr, Richter, N. J. & Merrick, W. C. Biochemical and kinetic characterization of the RNA helicase activity of eukaryotic initiation factor 4A. *J. Biol. Chem.* **274**, 12236–12244 (1999).
20. Neef, R. *et al.* Phosphorylation of mitotic kinesin-like protein 2 by polo-like kinase 1 is required for cytokinesis. *J. Cell Biol.* **162**, 863–875 (2003).
21. Liu, X., Zhou, T., Kuriyama, R. & Erikson, R. L. Molecular interactions of Polo-like-kinase 1 with the mitotic kinesin-like protein CHO1/MKLP-1. *J. Cell Sci.* **117**, 3233–3246 (2004).
22. Beretta, L., Gingras, A. C., Svitkin, Y. V., Hall, M. N. & Sonenberg, N. Rapamycin blocks the phosphorylation of 4E-BP1 and inhibits cap-dependent initiation of translation. *EMBO J.* **15**, 658–664 (1996).
23. Li, T., Inoue, A., Lahti, J. M. & Kidd, V. J. Failure to proliferate and mitotic arrest of CDK11$^{p110/p58}$-null mutant mice at the blastocyst stage of embryonic cell development. *Mol. Cell. Biol.* **24**, 3188–3197 (2004).
24. Petretti, C. *et al.* The PITSLRE/CDK11p58 protein kinase promotes centrosome maturation and bipolar spindle formation. *EMBO Rep.* **7**, 418–424 (2006).
25. Cornelis, S. *et al.* Identification and characterization of a novel cell cycle-regulated internal ribosome entry site. *Mol. Cell* **5**, 597–605 (2000).
26. Shi, Q. & King, R. W. Chromosome nondisjunction yields tetraploid rather than aneuploid cells in human cell lines. *Nature* **437**, 1038–1042 (2005).
27. Fujiwara, T. *et al.* Cytokinesis failure generating tetraploids promotes tumorigenesis in p53-null cells. *Nature* **437**, 1043–1047 (2005).
28. Brunet, A. *et al.* 14-3-3 transits to the nucleus and participates in dynamic nucleocytoplasmic transport. *J. Cell Biol.* **156**, 817–828 (2002).
29. Wilker, E. W., Grant, R. A., Artim, S. C. & Yaffe, M. B. A structural basis for 14-3-3σ functional specificity. *J. Biol. Chem.* **280**, 18891–18898 (2005).
30. Rubinson, D. A. *et al.* A lentivirus-based system to functionally silence genes in primary mammalian cells, stem cells and transgenic mice by RNA interference. *Nature Genet.* **33**, 401–406 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank P. Stern, D. Lowery and W. Merrick for reagents and technical assistance. This work was supported by postdoctoral fellowships from the Anna Fuller Fund and the NIH to E.W.W., an EMBO long-term fellowship to M.A.T.M.v.V., the David H. Koch Cancer Research Fund, NIH grants, and a Burroughs-Wellcome Career Development Award to M.B.Y.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.B.Y. (myaffe@mit.edu).

LETTERS

A ubiquitin ligase transfers preformed polyubiquitin chains from a conjugating enzyme to a substrate

Wei Li[1], Daqi Tu[2], Axel T. Brunger[2] & Yihong Ye[1]

In eukaryotic cells, many short-lived proteins are conjugated with Lys 48-linked ubiquitin chains and degraded by the proteasome[1]. Ubiquitination requires an activating enzyme (E1), a conjugating enzyme (E2) and a ligase (E3)[2]. Most ubiquitin ligases use either a HECT (homologous to E6-associated protein C terminus) or a RING (really interesting new gene) domain to catalyse polyubiquitination[3], but the mechanism of E3 catalysis is poorly defined[4]. Here we dissect this process using mouse Ube2g2 (E2; identical at the amino acid level to human Ube2g2) and human gp78 (E3), an endoplasmic reticulum (ER)-associated conjugating system essential for the degradation of misfolded ER proteins[5,6]. We demonstrate by expressing recombinant proteins in *Escherichia coli* that Ube2g2/gp78-mediated polyubiquitination involves preassembly of Lys 48-linked ubiquitin chains at the catalytic cysteine of Ube2g2. The growth of Ube2g2-anchored ubiquitin chains seems to be mediated by an aminolysis-based transfer reaction between two Ube2g2 molecules that each carries a ubiquitin moiety in its active site. Intriguingly, polyubiquitination of a substrate can be achieved by transferring preassembled ubiquitin chains from Ube2g2 to a lysine residue in a substrate.

To understand how Lys 48-linked ubiquitin chains are synthesized, we characterized a simple conjugating system consisting of a conserved E2 (Ube2g2) and gp78, a RING domain ubiquitin ligase. Because the amino-terminal transmembrane segments of gp78 are not required for its polyubiquitination function, we used its carboxy-terminal cytosolic domain (gp78c)[5,7]. Purified Ube2g2 and gp78c rapidly assembled Lys 48-linked polyubiquitin chains in a RING-dependent manner (Fig. 1a; Supplementary Fig. 1). Surprisingly, these chains were attached to a 20 kD protein via a thiol linkage, and were converted to unanchored species on addition of a reducing reagent, dithiothreitol (DTT) (Fig. 1a; Supplementary Fig. 2). Indeed, these ubiquitin chains could also be detected by immunoblotting with antibodies against Ube2g2, the only 20 kD protein in the reaction (Fig. 1b; Supplementary Fig. 3). The Ube2g2-linked ubiquitin chains were similarly assembled by full-length gp78 purified from mammalian cells (Supplementary Fig. 4). These chains were attached to the catalytic cysteine (Cys 89) of Ube2g2, because mutating the Cys 89, but not the other two cysteines in Ube2g2, abolished chain assembly (Fig. 1c; data not shown). The formation of Ube2g2-linked ubiquitin chains also required a conserved acidic loop of Ube2g2, which was located near its catalytic cysteine[8]. A loop mutant (2D2E) containing substitutions of four acidic residues by alanine failed to polymerize ubiquitin chains on itself (Fig. 1c). Because a similar loop has been found to be essential for the polyubiquitination function of a cell cycle specific E2, Cdc34 (ref. 9), the loop-dependent assembly of Ube2g2-anchored ubiquitin chains may be important for Ube2g2 and gp78-mediated polyubiquitination.

To understand how gp78c assembles ubiquitin chains on Ube2g2, we dissected the simplest step of this reaction—the formation of a di-ubiquitin on Ube2g2. One possible mechanism to generate Ube2g2-linked di-ubiquitin is by an aminolysis-based transfer reaction between two Ube2g2~ubiquitins, in which the Lys 48 in one Ube2g2-bound ubiquitin molecule attacks a donor ubiquitin that is conjugated to a neighbouring Ube2g2 (Fig. 2a). To test this idea, Ube2g2 was charged with either untagged wild-type ubiquitin or Flag-tagged K48R ubiquitin mutant (F–UbK48R). The reactions were quenched to block recharging of Ube2g2 (ref. 9). The two charging reactions were combined, and incubated in the presence of gp78c (Fig. 2b). If a transfer reaction between two Ube2g2~ubiquitins was feasible, addition of Ube2g2~wild-type-ubiquitin to Ube2g2~F–UbK48R should lead to the formation of Ube2g2-linked hybrid di-ubiquitin consisting of one wild-type ubiquitin and one F–UbK48R mutant, because Ube2g2~wild-type-ubiquitin could serve as an acceptor to receive the donor F–UbK48R (Fig. 2b). Indeed, addition of Ube2g2~wild-type-ubiquitin, but not a charging reaction containing either no Ube2g2 (mock) or the Ube2g2 C89A mutant, to Ube2g2~F–UbK48R resulted in the formation of Ube2g2-anchored hybrid di-ubiquitin (Fig. 2b, c). When more Ube2g2~F–UbK48R and Ube2g2~wild-type-ubiquitin molecules were used (by increasing the precharging time), a small fraction of Ube2g2 conjugated with two wild-type ubiquitins and one F–UbK48R was formed (Fig. 2c, lanes 2, 3). These results validate the proposed transfer reaction. They also indicate that longer ubiquitin chains may be formed on Ube2g2 by multiple rounds of such transfer reactions.

Because the Ube2g2 loop mutant (2D2E) failed to assemble ubiquitin chains on itself (Fig. 1c), the ubiquitin transfer reaction probably involves the acid loop. Indeed, when Ube2g2 2D2E~ubiquitin was used as acceptor, the rate of F–UbK48R transfer from wild-type Ube2g2 was dramatically reduced (Fig. 2d, lanes 5–8 versus lanes 1–4). When the 2D2E mutant conjugated with F–UbK48R was used as donor, the transfer reaction was completely abolished (lanes 9–16). Together, these data suggest that the acidic loop of Ube2g2 has a critical role with the donor ubiquitin during the assembly of Ube2g2-linked ubiquitin chains.

To determine whether Ube2g2-linked ubiquitin chains are involved in substrate polyubiquitination, we used an ER-associated protein termed HERP because it is a short-lived protein that interacts with several E3 ligases including gp78 (refs 10–12). We purified the cytosolic domain of human HERP (HERPc) that contained the E3 binding site[13], and found that it was rapidly polyubiquitinated by Ube2g2 and gp78c (Supplementary Fig. 5). Polyubiquitination of HERPc required the gp78 RING domain (Fig. 3a) and the Lys 48 of ubiquitin (Supplementary Fig. 5). Mass spectrometry analysis identified Lys 61 in HERPc as the site of polyubiquitination (data not shown). Indeed, a HERPc K61R mutant failed to undergo

[1]Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA. [2]The Howard Hughes Medical Institute, Department of Molecular and Cellular Physiology, Neurology and Neurological Sciences, Structural Biology, and SSRL, Stanford University, Stanford, California 94305-5489, USA.

polyubiquitination by gp78c (Fig. 3b). Thus, gp78c and Ube2g2 conjugate a single Lys 48-linked ubiquitin chain to Lys 61 in HERPc. We next compared the rate of chain assembly on Ube2g2 with that of HERPc polyubiquitination (Supplementary Fig. 6). The formation of Ube2g2-linked chains slightly preceded the appearance of polyubiquitinated HERPc, raising the possibility that ubiquitin chains may be first assembled on Ube2g2 before being transferred to HERPc. Consistent with this idea, fewer Ube2g2-linked ubiquitin chains were generated when wild-type HERPc was present compared with those formed in the presence of HERPc K61R (Supplementary Fig. 7). Presumably, this was owing to the simultaneous transfer of ubiquitin chains from Ube2g2 to HERPc.

To test further the preassembly idea, we analysed HERPc polyubiquitination using the Ube2g2 loop mutant (2D2E) because it was defective in assembling Ube2g2-linked ubiquitin chains. Indeed, this mutant was able to conjugate one ubiquitin to HERPc, but was inactive in conjugating longer chains to HERPc (Fig. 3c). To obtain further evidence, we mutated the highly conserved Tyr 83 and His 94 of Ube2g2, respectively, because these amino acids seemed to communicate with the catalytic cysteine, and were located near the acid loop (Supplementary Fig. 8a–c). The Y83A mutant was able to assemble Ube2g2-linked ubiquitin chains, albeit with reduced activity; whereas the H94A mutant was almost completely inactive (Supplementary Fig. 8d). Consistent with our interpretation, the Y83A mutant exhibited similarly reduced activity towards ubiquitination of HERPc, whereas the H94A mutant was only capable of conjugating shorter ubiquitin oligomers ($n < 3$) to HERPc (Supplementary Fig. 8e). Furthermore, when similar mutations were introduced in the yeast homologue Ubc7, the mutant proteins were defective in rescuing the growth defect of a mutant yeast strain that lacked both the ER degradation and the unfolded protein response pathways (Δ*Ubc7* and Δ*Ire1* double deletion strain) (Supplementary Fig. 8f). Together, these data indicate that preassembly of Ube2g2-linked ubiquitin chains is essential for the function of Ube2g2 *in vitro* and *in vivo*.

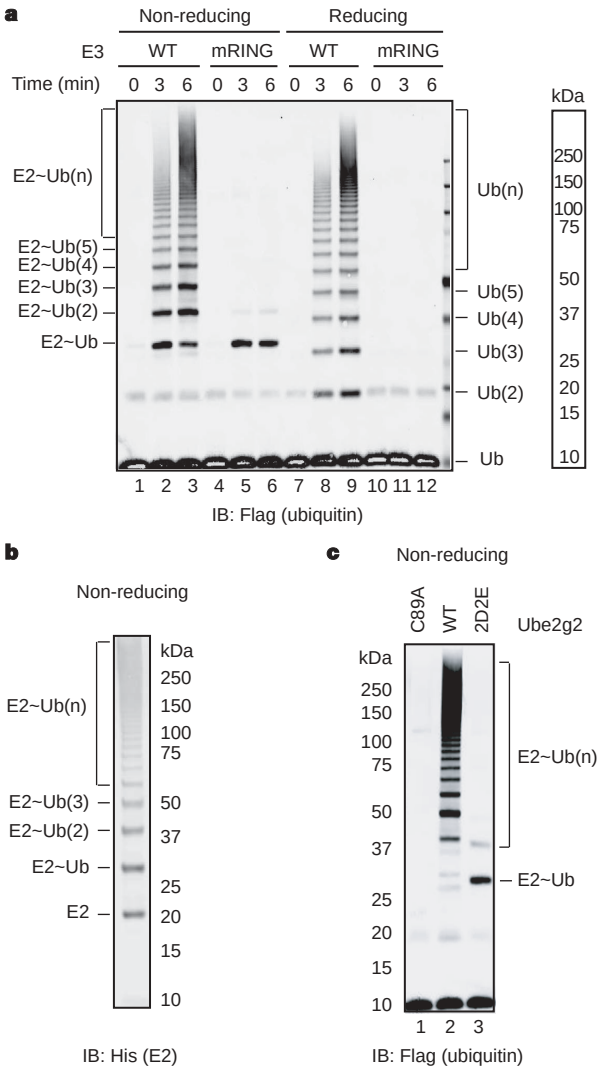


Figure 1 | Formation of Ube2g2-linked ubiquitin chains. a, gp78c (E3) assembles ubiquitin chains in a RING-dependent manner. Ubiquitination reactions performed with Flag-tagged ubiquitin in the presence of either wild-type (WT) gp78c or its RING mutant (mRING) were analysed by immunoblotting (IB) with anti-Flag antibodies. E2~Ub(n), E2-linked ubiquitin (Ub) chains consisting of a certain number (n) of ubiquitin molecules. Ub(n), unanchored ubiquitin chains. **b**, Ubiquitin chains are linked to Ube2g2 (E2). Ubiquitin chains assembled by Ube2g2 and gp78c were analysed by immunoblotting with anti-His antibodies. **c**, Formation of Ube2g2-linked ubiquitin chains requires the catalytic cysteine (C89) and the acidic loop of Ube2g2. Ubiquitination reactions were performed as in **a**, except that the indicated E2 variants were used.

To prove directly that Ube2g2-linked ubiquitin chains are transferable, we developed a two-step ubiquitination assay (Fig. 3d). First, Ube2g2-anchored ubiquitin chains were assembled in the absence of HERPc using a low concentration of ubiquitin. This ensured that most ubiquitin molecules were incorporated into Ube2g2-linked chains (Fig. 3d, lane 1; Supplementary Fig. 9). The reaction was stopped by dialysis, which removed ATP. Ube2g2-linked ubiquitin chains were then incubated with either HERPc or HERPc K61R. After incubation, ubiquitin chains were indeed transferred to HERPc, but not to HERPc K61R (Fig. 3d, lanes 2–7). These chains could not be generated from Ube2g2 that carried only one ubiquitin moiety because few such molecules were present in the preassembly reaction. In addition, incubation of HERPc with Ube2g2 precharged with one ubiquitin moiety formed only shorter ubiquitin chains on HERPc (Supplementary Fig. 9). Presumably, removal of ATP prevented recharging of Ube2g2, which rendered the chain assembly process less efficient. Together, these data demonstrate that ubiquitin chains assembled on Ube2g2 can be transferred 'en bloc' to a lysine residue in a substrate. They also indicate that the processivity of gp78-dependent polyubiquitination (the number of ubiquitin molecules transferred in one binding event) may be regulated by controlling the rate of chain transfer to substrate versus that of chain assembly on Ube2g2.

To prove further the 'chain transfer' idea, we tested whether chemically synthesized ubiquitin polymers can be directly transferred to HERPc. It was shown previously that ubiquitin chains consisting of up to four ubiquitins (Ub4) could be charged onto Ubc4 and Ubc7 (ref. 14). We confirmed that chemically synthesized Ub2, Ub3 and Ub4 could also be conjugated to the catalytic cysteine of Ube2g2 (Supplementary Fig. 10). When ubiquitination of HERPc was performed with Ub2, Ub3, or Ub4, a set of ubiquitinated HERPc species was generated, apparently by transfer of the corresponding ubiquitin oligomers from Ube2g2 to HERPc (Fig. 4a). The transfer of preassembled ubiquitin chains to substrates still required the gp78 RING domain, because the RING mutant gp78c was unable to conjugate Ub2 and Ub3 to HERPc (Fig. 4b). Together, these data confirm that ubiquitin chains can be transferred 'en bloc' from a conjugating enzyme to a substrate in the presence of a functional RING ligase.

To see whether assembly of Ube2g2-linked ubiquitin chains occurs *in vivo*, we expressed Flag-tagged Ube2g2 together with HA-tagged ubiquitin[15] in 293T cells. Because Ube2g2-linked ubiquitin chains are probably quite labile in cells (either as a result of simultaneous transfer to substrates or rapid turn-over by the proteasome), cells were treated with a proteasome inhibitor to accumulate such species. As expected, only a small fraction of Ube2g2 carried polyubiquitin conjugates (Fig. 4c, lanes 2, 5), and these conjugates could be removed from Ube2g2 by treatment with DTT (lanes 3). Thus, ubiquitin

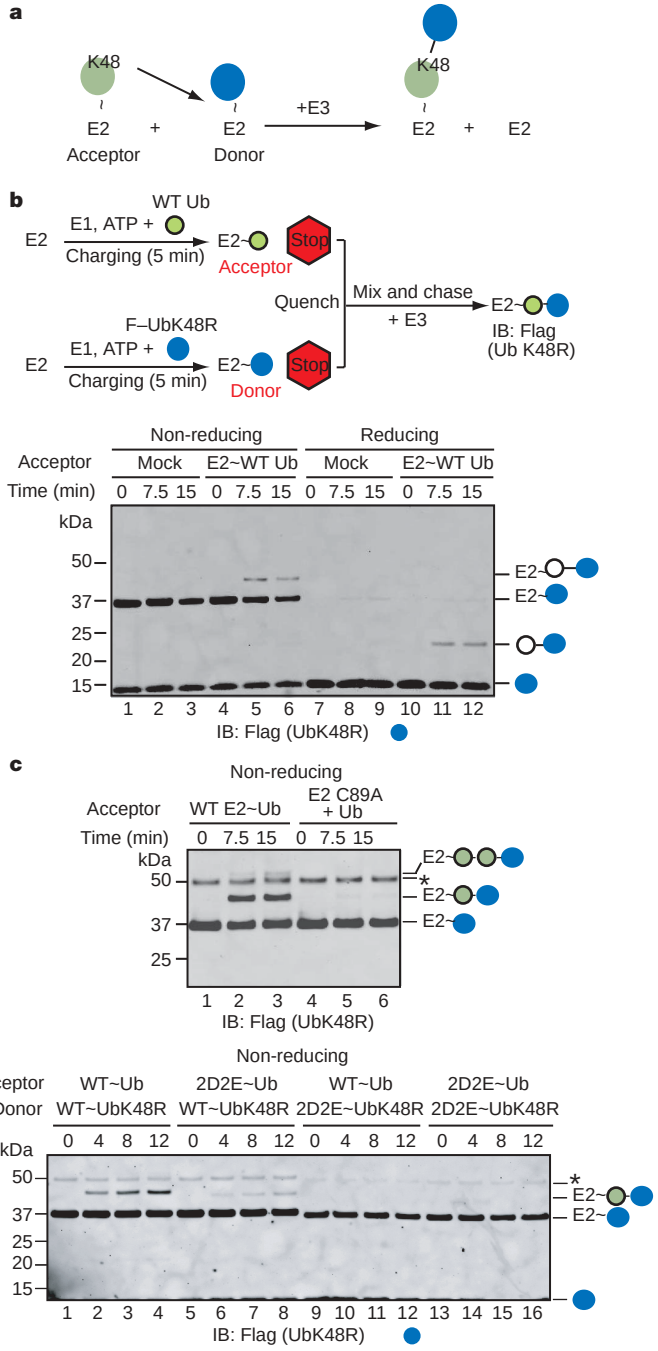


Figure 2 | Formation of Ube2g2-linked di-ubiquitin by an aminolysis-based transfer reaction. **a**, Outline of the proposed transfer reaction between two E2~Ubs. (K48, Lys48; ~ indicates thioester bond). **b**, Formation of Ube2g2-linked hybrid di-ubiquitin. Ube2g2 (E2) charged with Flag-tagged UbK48R (F–UbK48R, blue circle) was treated with EDTA/NEM (stop), and mixed with equal molar concentration of EDTA/NEM-quenched Ube2g2 precharged with untagged wild-type ubiquitin (green circle). The chase reaction was conducted in the presence of gp78c (E3). The reaction was analysed by immunoblotting with anti-Flag antibodies to detect products containing F–UbK48R. Where indicated, Ube2g2 charged with F–UbK48R was mixed with a mock charging reaction containing no Ube2g2 (mock). **c**, As **b**, except that the charging of Ube2g2 with either F–UbK48R or wild-type ubiquitin was conducted for a longer period of time (15 min). Where indicated, charging reaction containing the Ube2g2 C89A mutant (E2 C89A + Ub) was used as the acceptor. The asterisk indicates Ube2g2 conjugated with two F–UbK48Rs that was probably generated by transfer of one F–UbK48R molecule to a lysine residue in Ube2g2 during the charging reaction. **d**, As in **c**, except that the indicated E2s precharged with the indicated ubiquitin variants were used.

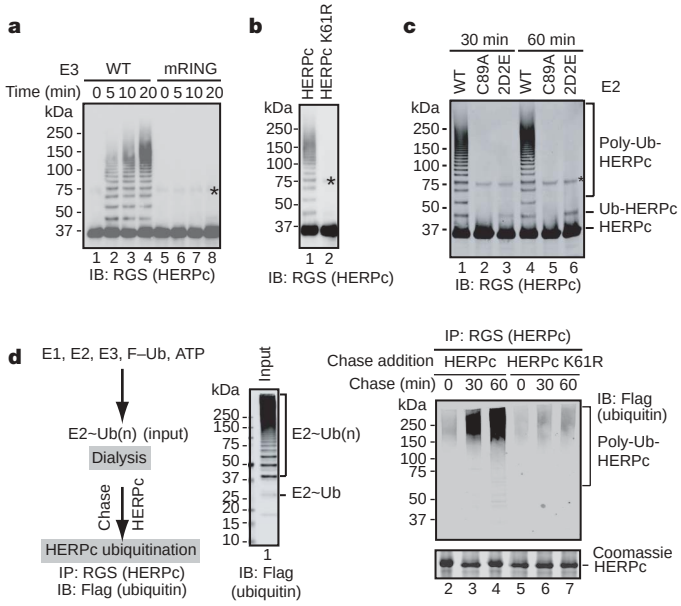


Figure 3 | Transfer of ubiquitin chains from Ube2g2 to HERPc. **a**, Ubiquitination of HERPc by gp78c. Ubiquitination of RGS-His-tagged HERPc was performed in the presence of either wild-type gp78c (E3) or its RING mutant (mRING). The asterisk marks a fraction of aggregated HERPc species. Unless otherwise specified, asterisks in the following panels refer to the same. **b**, Lys 61 of HERPc is required for its polyubiquitination. HERPc or a HERPc mutant bearing a Lys 61 to Arg (K61R) substitution was subjected to ubiquitination as in **a**. **c**, Polyubiquitination of HERPc requires the acidic loop of Ube2g2. Ubiquitination of HERPc was performed with the indicated Ube2g2 (E2) variants. **d**, Ube2g2-linked ubiquitin chains (E2~Ub(n)) were first synthesized in the presence of E1, E2, E3, Flag-tagged ubiquitin (F–Ub) and ATP. A fraction of the reaction was directly analysed by immunoblotting with anti-Flag antibodies (input), whereas the rest of the reaction was subjected to dialysis to remove ATP. The dialysed product was incubated with either HERPc or the HERPc K61R mutant for the indicated time points (chase). Ubiquitinated HERPc was detected by immunoblotting the immunoprecipitated HERPc with anti-Flag antibodies. A fraction of the immunoprecipitated HERPc was analysed by SDS–PAGE and Coomassie blue staining.

chains can be assembled on Ube2g2 in a thiol-dependent manner in cells.

In this report, we demonstrate that ubiquitin chains can be pre-assembled on the catalytic cysteine of an E2 (Ube2g2) before being transferred to a substrate (Fig. 4d). The proposed 'preassembly' model is distinct from the generally presumed 'sequential addition' model, in which ubiquitin molecules are added one at a time, first to a lysine residue in a substrate, and then to the Lys 48 in the distal ubiquitin of a growing chain. Our data indicate that Ube2g2-linked ubiquitin chains are assembled by multiple rounds of transfer reactions between two Ube2g2s bound by ubiquitin. To efficiently transfer ubiquitin, Ube2g2 may need to form a dimer. Indeed, a small fraction (~5%) of the purified Ube2g2 does exist as dimer. Moreover, gp78 forms homo-oligomers using its cytosolic domain in cells (Supplementary Fig. 11). Thus, gp78 may promote the efficient transfer of ubiquitin between two Ube2g2~ubiquitins by facilitating their dimerization. Dimerization of E2 has been observed for other conjugating enzymes including Cdc34 (refs 16–19). In fact, the activity of Cdc34 is significantly enhanced under conditions that favour its dimerization[20]. These observations together with the fact that Cdc34 also contains a similar acidic loop indicate that it may operate by a similar mechanism. Intriguingly, like Ube2g2, many E2 enzymes are able to assemble ubiquitin chains independent of substrates[21–26]. Further characterization of these reactions may reveal additional examples of thiol-linked ubiquitin chains on E2 enzymes.

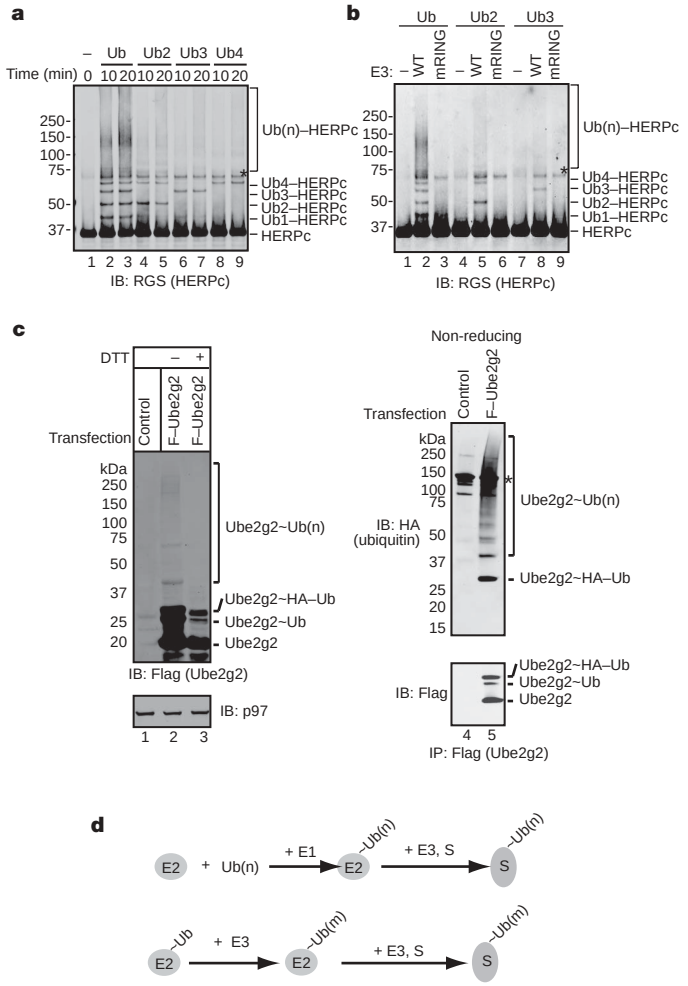


Figure 4 | Ubiquitination of HERPc by chemically synthesized ubiquitin oligomers. a, Ubiquitin oligomers can be directly transferred from Ube2g2 to HERPc. Ubiquitination of HERPc performed in the presence of ubiquitin or the indicated ubiquitin oligomers was analysed by immunoblotting (IB) with anti-RGS-His antibodies. **b**, Transfer of ubiquitin oligomers from Ube2g2 to HERPc requires the gp78c RING domain. Ubiquitination of HERPc was conducted with the indicated ubiquitin oligomers in the absence of (−) or in the presence of wild-type gp78c (E3) or its RING mutant (mRING). **c**, Formation of thiol-linked ubiquitin chains on Ube2g2 in cells. 293T cells transfected with a construct expressing haemagglutinin (HA)-tagged ubiquitin together with either an empty vector (control) or a plasmid expressing Flag-tagged Ube2g2 were treated with a proteasome inhibitor MG132. Cell extracts were analysed by immunoblotting with the indicated antibodies. A fraction of the extracts was subjected to immunoprecipitation with anti-Flag antibodies before immunoblotting analysis. Asterisk, IgG. **d**, A model for Ube2g2 (E2)- and gp78 (E3)-mediated polyubiquitination. Upper panel, ubiquitin chains present in cells[13] can be directly charged onto an E2 and then transferred to a substrate (S) in the presence of an E3. Alternatively (lower panel), longer chains can be assembled on E2 before being transferred to substrates. Ub(m), Ubiquitin chains containing a certain number (m) of ubiquitin molecules.

METHODS

Detailed methods are presented in Supplementary Information.

Antibodies and proteins. Anti-Flag and anti-RGS-His antibodies (antibodies recognizing Arg-Gly-Ser plus 4 consecutive His residues) were from Sigma and Qiagen, respectively.

Flag-tagged ubiquitin, Myc-tagged ubiquitin, untagged bovine ubiquitin, untagged UbK48R, methylated ubiquitin, E1, and chemically synthesized ubiquitin oligomers were purchased from Boston Biochem.

Ubiquitination assay. Unless specified in the figure legends, ubiquitination experiments were performed as described previously[7]. Briefly, E1 (60 nM), Ube2g2 (200 nM), gp78c (300 nM) were incubated with Flag-tagged ubiquitin (10 μM) at 37 °C in buffers containing 25 mM Tris HCl, pH 7.4, 2 mM magnesium/ATP, and 0.1 mM DTT. For reducing conditions, samples were treated with 100 mM DTT before SDS–polyacrylamide gel electrophoresis analyses. Ubiquitination of HERPc was conducted using the same conditions described above with the addition of HERPc (500 nM). Ubiquitinated HERPc was detected by immunoblotting with anti-RGS-His antibody.

Ubiquitin transfer assay. To monitor the transfer of ubiquitin between two Ube2g2s, Ube2g2 (1 μM) was incubated with E1 (60 nM), Flag-tagged UbK48R or untagged wild-type ubiquitin (1 μM) at 25 °C in buffers containing 25 mM Tris HCl, pH 7.4, and 2 mM magnesium/ATP. The reactions were treated with 50 mM EDTA and 10 mM NEM for 15 min at 25 °C to prevent further rounds of charging. The two reactions were then combined and incubated (chase) at 37 °C in the presence of gp78c (300 nM). The reaction was stopped by Laemmli buffer and analysed by immunoblotting with anti-Flag antibodies.

Received 21 November; accepted 18 December 2006.
Published online 18 February 2007.

1. Chau, V. *et al.* A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science* **243,** 1576–1583 (1989).
2. Scheffner, M., Nuber, U. & Huibregtse, J. M. Protein ubiquitination involving an E1–E2–E3 enzyme ubiquitin thioester cascade. *Nature* **373,** 81–83 (1995).
3. Pickart, C. M. & Eddins, M. J. Ubiquitin: structures, functions, mechanisms. *Biochim. Biophys. Acta* **1695,** 55–72 (2004).
4. Hochstrasser, M. Lingering mysteries of ubiquitin-chain assembly. *Cell* **124,** 27–34 (2006).
5. Fang, S. *et al.* The tumor autocrine motility factor receptor, gp78, is a ubiquitin protein ligase implicated in degradation from the endoplasmic reticulum. *Proc. Natl Acad. Sci. USA* **98,** 14422–14427 (2001).
6. Chen, B. *et al.* The activity of a human endoplasmic reticulum-associated degradation E3, gp78, requires its Cue domain, RING finger, and an E2-binding site. *Proc. Natl Acad. Sci. USA* **103,** 341–346 (2006).
7. Ye, Y., Meyer, H. H. & Rapoport, T. A. Function of the p97–Ufd1–Npl4 complex in retrotranslocation from the ER to the cytosol: dual recognition of nonubiquitinated polypeptide segments and polyubiquitin chains. *J. Cell Biol.* **162,** 71–84 (2003).
8. Arai, R. *et al.* Structure of human ubiquitin-conjugating enzyme E2 G2 (UBE2G2/UBC7). *Acta Crystallogr. F* **62,** 330–334 (2006).
9. Petroski, M. D. & Deshaies, R. J. Mechanism of lysine 48-linked ubiquitin-chain synthesis by the cullin-RING ubiquitin-ligase complex SCF–Cdc34. *Cell* **123,** 1107–1120 (2005).
10. Sai, X. *et al.* The ubiquitin-like domain of Herp is involved in Herp degradation, but not necessary for its enhancement of amyloid β-protein generation. *FEBS Lett.* **553,** 151–156 (2003).
11. Hori, O. *et al.* Role of Herp in the endoplasmic reticulum stress response. *Genes Cells* **9,** 457–469 (2004).
12. Schulze, A. *et al.* The ubiquitin-domain protein HERP forms a complex with components of the endoplasmic reticulum associated degradation pathway. *J. Mol. Biol.* **354,** 1021–1027 (2005).
13. Kokame, K., Agarwala, K. L., Kato, H. & Miyata, T. Herp, a new ubiquitin-like membrane protein induced by endoplasmic reticulum stress. *J. Biol. Chem.* **275,** 32846–32853 (2000).
14. van Nocker, S. & Vierstra, R. D. Multiubiquitin chains linked through lysine 48 are abundant *in vivo* and are competent intermediates in the ubiquitin proteolytic pathway. *J. Biol. Chem.* **268,** 24766–24773 (1993).
15. Vong, Q. P., Cao, K., Li, H. Y., Iglesias, P. A. & Zheng, Y. Chromosome alignment and segregation regulated by ubiquitination of survivin. *Science* **310,** 1499–1504 (2005).
16. Silver, E. T., Gwozd, T. J., Ptak, C., Goebl, M. & Ellison, M. J. A chimeric ubiquitin conjugating enzyme that combines the cell cycle properties of CDC34 (UBC3) and the DNA repair properties of RAD6 (UBC2): implications for the structure, function and evolution of the E2s. *EMBO J.* **11,** 3091–3098 (1992).
17. Chen, P., Johnson, P., Sommer, T., Jentsch, S. & Hochstrasser, M. Multiple ubiquitin-conjugating enzymes participate in the *in vivo* degradation of the yeast MATα2 repressor. *Cell* **74,** 357–369 (1993).
18. Varelas, X., Ptak, C. & Ellison, M. J. Cdc34 self-association is facilitated by ubiquitin thiolester formation and is required for its catalytic activity. *Mol. Cell. Biol.* **23,** 5388–5400 (2003).
19. Eddins, M. J., Carlile, C. M., Gomez, K. M., Pickart, C. M. & Wolberger, C. Mms2–Ubc13 covalently bound to ubiquitin reveals the structural basis of linkage-specific polyubiquitin chain formation. *Nature Struct. Mol. Biol.* **13,** 915–920 (2006).
20. Gazdoiu, S. *et al.* Proximity-induced activation of human Cdc34 through heterologous dimerization. *Proc. Natl Acad. Sci. USA* **102,** 15053–15058 (2005).
21. Lorick, K. L. *et al.* RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination. *Proc. Natl Acad. Sci. USA* **96,** 11364–11369 (1999).

22. Joazeiro, C. A. *et al.* The tyrosine kinase negative regulator c-Cbl as a RING-type, E2-dependent ubiquitin-protein ligase. *Science* **286,** 309–312 (1999).
23. Tan, P. *et al.* Recruitment of a ROC1–CUL1 ubiquitin ligase by Skp1 and HOS to catalyze the ubiquitination of IκBα. *Mol. Cell* **3,** 527–533 (1999).
24. Huang, H. *et al.* The inhibitor of apoptosis, cIAP2, functions as a ubiquitin-protein ligase and promotes *in vitro* monoubiquitination of caspases 3 and 7. *J. Biol. Chem.* **275,** 26661–26664 (2000).
25. Bays, N. W., Gardner, R. G., Seelig, L. P., Joazeiro, C. A. & Hampton, R. Y. Hrd1p/Der3p is a membrane-anchored ubiquitin ligase required for ER-associated degradation. *Nature Cell Biol.* **3,** 24–29 (2001).
26. Lu, Z., Xu, S., Joazeiro, C., Cobb, M. H. & Hunter, T. The PHD domain of MEKK1 acts as an E3 ubiquitin ligase and mediates ubiquitination and degradation of ERK1/2. *Mol. Cell* **9,** 945–956 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank T. Rapoport, M. Gellert and M. Krause for critical reading of the manuscript; A. Weissman, Y. Zheng and M. Seeger for plasmids; P. Carvalho for the Δ*Ubc7* Δ*Ire1* strain; M. Hochstrasser for communicating results before publication; and the Taplin Biological Mass Spectrometry Facility at Harvard Medical School for mass spectrometry analysis. This work was supported by funding from the NIDDK intramural research program at the National Institutes of Health.

Author Contributions W.L. performed most of the experiments. D.T. and Y.Y. performed some of the experiments. D.T. and A.T.B. provided intellectual input. W.L. and Y.Y. designed the experiments. Y.Y. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to Y.Y. (yihongy@mail.nih.gov).

LETTERS

Structure and function of the histone chaperone CIA/ASF1 complexed with histones H3 and H4

Ryo Natsume[1]*, Masamitsu Eitoku[2]*, Yusuke Akai[1], Norihiko Sano[2], Masami Horikoshi[2,3] & Toshiya Senda[4]

CIA (CCG1-interacting factor A)/ASF1, which is the most conserved histone chaperone among the eukaryotes, was genetically identified as a factor for an anti-silencing function (Asf1)[1] by yeast genetic screening. Shortly after that, the CIA–histone-H3–H4 complex was isolated from *Drosophila* as a histone chaperone CAF-1 stimulator[2]. Human CIA-I/II (ASF1a/b) was identified as a histone chaperone that interacts with the bromodomain—an acetylated-histone-recognizing domain—of CCG1, in the general transcription initiation factor TFIID[3–5]. Intensive studies have revealed that CIA/ASF1 mediates nucleosome assembly by forming a complex with another histone chaperone in human cells[6] and yeast[7], and is involved in DNA replication[1,2], transcription[4,8–10], DNA repair[1,2,11,12] and silencing/anti-silencing[1,2,8,13–15] in yeast. CIA/ASF1 was shown as a major storage chaperone for soluble histones in proliferating human cells[6,16]. Despite all these biochemical and biological functional analyses, the structure–function relationship of the nucleosome assembly/disassembly activity of CIA/ASF1 has remained elusive. Here we report the crystal structure, at 2.7 Å resolution, of CIA-I in complex with histones H3 and H4. The structure shows the histone H3–H4 dimer's mutually exclusive interactions with another histone H3–H4 dimer and CIA-I. The carboxy-terminal β-strand of histone H4 changes its partner from the β-strand in histone H2A to that of CIA-I through large conformational change. *In vitro* functional analysis demonstrated that CIA-I has a histone H3–H4 tetramer-disrupting activity. Mutants with weak histone H3–H4 dimer binding activity showed critical functional effects on cellular processes related to transcription. The histone H3–H4 tetramer-disrupting activity of CIA/ASF1 and the crystal structure of the CIA/ASF1–histone-H3–H4 dimer complex should give insights into mechanisms of both nucleosome assembly/disassembly and nucleosome semi-conservative replication.

The CIA family in multicellular organisms consists of two members, CIA-I and CIA-II, which correspond to ASF1a and ASF1b, respectively[3,5,17]. Although the two CIAs are similar in primary sequences, the messenger RNA distributions are different. CIA-I mRNA is widely distributed throughout the body, whereas CIA-II mRNA is highly cell-specific to the thymus, testis, small intestine and colon[5]. We used CIA-I for our structural and functional analyses to elucidate the basic and general roles of CIA activity. The crystal structure of human CIA-I(172) (residues 1–172) in complex with core histones H3 (*Xenopus laevis* H3.1) and H4 (*X. laevis* H4) (the CIA-I–histone-H3–H4 complex) was determined at 2.7 Å resolution (Supplementary Fig. 1; Supplementary Table 1). Because human and *Xenopus* histones H3 and H4 differ by only one amino acid in total, which is not located at the interacting surface of the complex, our structure can be regarded as a homologous system.

In the crystal, CIA-I forms a complex with one histone H3–H4 dimer (Fig. 1a). The stoichiometry of the complex is consistent with earlier biochemical results[6,16,18]. It is striking that CIA-I and the histone H3–H4 residues that are important for interaction with other factors are mostly exposed to the solvent in the present complex. First, residues that are different between CIA-I and -II are exposed, indicating that CIA-I and -II seem to interact with the histone H3–H4 dimer essentially in the same manner, and CIA-variant-specific factors could distinguish CIA-I and -II in the complex. Second, possible modification sites of not only the histone tail region[19] but also the histone core domains[20], namely Lys 79(H3), Lys 115(H3), Thr 118(H3), Lys 122(H3), Lys 31(H4), Ser 47(H4), Lys 59(H4), Lys 77(H4), Lys 79(H4), Lys 91(H4) and Arg 92(H4), are exposed (Fig. 1b; Supplementary Fig. 2). Although Lys 115(H3), Lys 122(H3), Lys 91(H4) and Arg 92(H4) are close to the CIA-I interaction regions, they are still exposed. The other residues are fully exposed and accessible to other chromatin enzymes/factors. Third, because residues that differ between histones H3.1 and H3.3 are exposed, it is possible that histone-variant-specific factors like CAF-1 and HIRA recognize them (Fig. 1b).

The histone H3–H4 dimer has two binding sites for CIA-I, one composed of helix α3 and the C-terminal half of helix α2 of histone H3 (residues 106–131), and the other composed of the C-terminal fragment of histone H4 (residues 95–100). The two binding sites of the histone H3–H4 dimer are also used for histone octamer formation[21] (Fig. 1c–f). It is striking that the C-terminal β-strand of histone H4 changes its partner from the β-strand in histone H2A to that of CIA-I through a large conformational change (Fig. 1e, f). The 'mutually exclusive binding mode' of the histone H3–H4 dimer suggests that CIA-I has histone H3–H4 tetramer-disrupting activity to form the CIA-I–histone-H3–H4 complex. It is of note that the crystal of the CIA-I–histone-H3–H4 complex appeared in a mixture of a histone H3–H4 tetramer and human CIA-I(172). In addition, the GST pull-down assay[3] suggested their interaction is stoichiometric (Fig. 1g). The molar ratio between human CIA-I(155) and the histone H3–H4 dimer in the saturated region was calculated as approximately 1:1 from densitometry.

To confirm the histone H3–H4 tetramer-disrupting activity of CIA-I, the subunit rearrangement between the histone H3–H4 tetramer and human CIA-I(172) was analysed in solution using a static light-scattering technique (Fig. 1h). The histone H3–H4 complex existed as a tetramer in solution with an estimated molecular weight of 57,290 ± 760 Da (theoretical value of 53,012 Da). The addition of a nearly equimolar amount of human CIA-I(172) to the solution caused the histone H3–H4 tetramer to disappear. The emerging new molecular species had a molecular weight of 49,600 ± 1,580 Da, which closely approximates the human CIA-I(172)–histone-H3–H4

[1]Japan Biological Information Research Centre (JBIRC), Japan Biological Informatics Consortium (JBIC), 2-42 Aomi, Koto-ku, Tokyo 135-0064, Japan. [2]Laboratory of Developmental Biology, Institute of Molecular and Cellular Biosciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan. [3]Horikoshi Gene Selector Project, Exploratory Research for Advanced Technology (ERATO), Japan Science and Technology Agency (JST), 5-9-6 Tokodai, Tsukuba, Ibaraki 300-2635, Japan. [4]Biological Information Research Centre (BIRC), National Institute of Advanced Industrial Science and Technology (AIST), 2-42 Aomi, Koto-ku, Tokyo 135-0064, Japan.
*These authors contributed equally to this work.

complex (theoretical value = 46,318 Da), and contains human CIA-I(172) and the histones H3 and H4 (Supplementary Fig. 3). Thus, we concluded that CIA-I disrupted the histone H3–H4 tetramer to form the CIA-I–histone-H3–H4 complex.

Next, we examined the histone-interacting sites on CIA-I. There are two surfaces on CIA-I that interact with the histone H3–H4 dimer. The primary binding site (924 Å^2) is on the front β-sheet, and the secondary binding site (434 Å^2) is located at the side edge of the two-layered β-sheets (β1- and β10-strands) (Fig. 1a). Although the binding sites have been partially revealed through the analysis of a two-hybrid assay[3] and chemical shifts in the NMR spectroscopy[22], the present structure shows details of the interaction between CIA-I and the histone H3–H4 dimer.

The primary binding site interacts with helix α3 and the C-terminal half of helix α2 of histone H3 (Fig. 1a). This polypeptide region undergoes a conformational change on complex formation to avoid close contact with the front β-sheet of CIA-I (Supplementary Fig. 4; Supplementary Table 2). Residues in the primary binding site form a hydrophobic concave groove, which accommodates Leu 126(H3), Arg 129(H3) and Ile 130(H3) (Fig. 2a, b). In addition, the primary binding site and its interacting surface on histone H3 have complementary charge distributions (Supplementary Fig. 5).

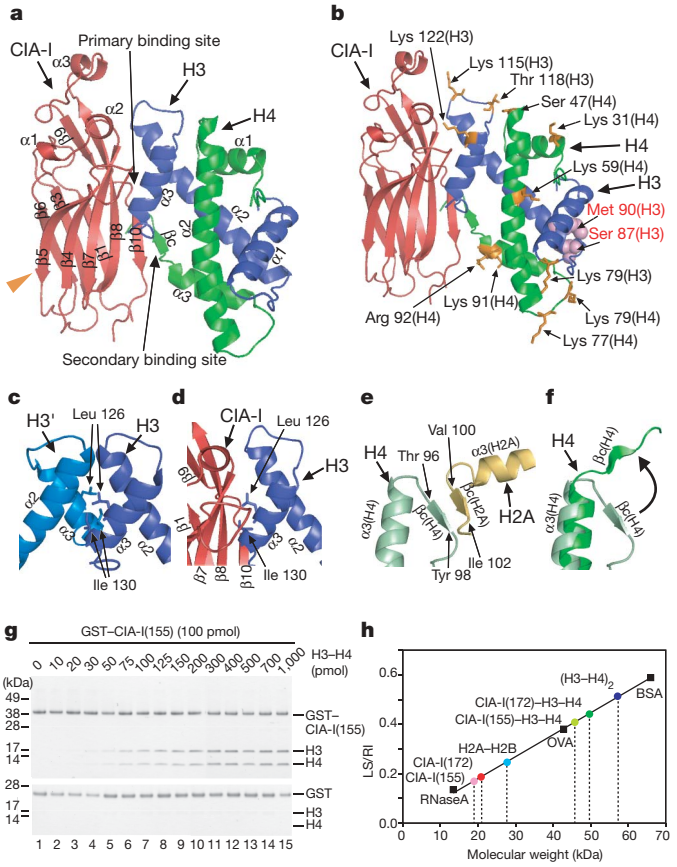


Figure 1 | Overall structure and biochemical analysis of the CIA-I–histone-H3–H4 complex. **a**, The overall structure of the CIA-I–histone-H3–H4 complex. CIA-I and histones H3 and H4 are shown in red, blue, and green, respectively. An orange triangle indicates the HIRA (histone regulatory homologue A) binding site[28]. The molecular graphics were prepared by PyMOL[29]. **b**, Possible modification sites (stick model in orange) and the different residues of human histone H3.1 and H3.3 in the CIA-I–histone-H3–H4 complex (sphere model in pink, only exposed residues are displayed). **c**, The histone H3–H3' interaction in the nucleosome core[21]. **d**, CIA-I–histone H3 interaction in the CIA-I–histone-H3–H4 complex. **e**, Carboxy-terminal residues (Thr 96–Tyr 98) of histone H4 (green) form a parallel β-sheet with a β-strand from H2A (yellow) in the nucleosome core[21]. **f**, The C-terminal fragment of histone H4 undergoes a conformational change on the CIA-I–histone-H3–H4 complex formation through rotations of main-chain ψ angles of Gly 94 and Arg 95 by 95° and 75°, respectively. Histone H4s in the present structure and the nucleosome core are shown in green and light green, respectively. This conformational change and formation of a new anti-parallel β-sheet might facilitate the histone H3–H4 tetramer-disruption. The large effect (tetramer disruption) originating from a small interaction (formation of β-sheet) is similar to the essence of Japanese Judo (Yawara): 'softness tames toughness (ju yoku go wo seisu)'. Following the spirit of Judo, we designated the mechanism as the 'Yawara split'. It is intriguing to note that most histone chaperones have a β-rich structure. **g**, Pull-down assay with GST-tagged human CIA-I(155) or GST (control), showing the stoichiometric interaction between CIA and the histone H3–H4 complex. **h**, Determination of the molecular weight of human CIA, histone dimer, tetramer, and complexes by the static light-scattering method. LS and RI represent the intensity of static light scattering and the refractive index, respectively.

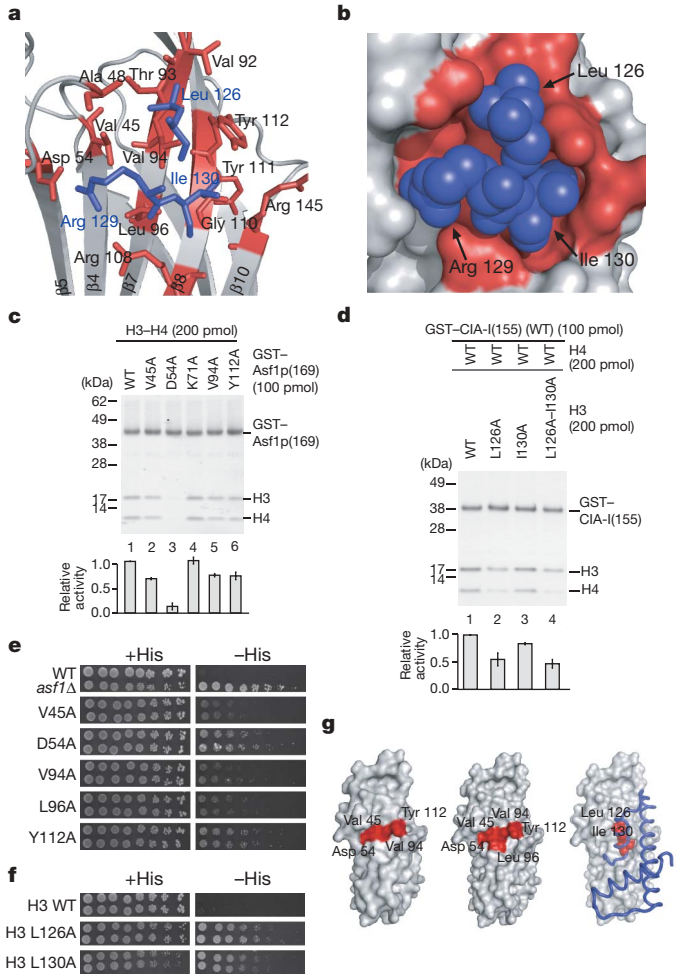


Figure 2 | Physical and functional interaction between CIA and histone H3. **a**, Interaction between the residues in the hydrophobic concave groove (red) of CIA-I and histone H3 (blue). **b**, The hydrophobic concave groove (red) in the primary binding site of CIA-I accommodates Leu 126(H3), Arg 129(H3) and Ile 130(H3) (blue). **c**, GST pull-down assay for the histone H3–H4 complex by GST-tagged wild-type or mutant yeast Asf1p/Cia1p(169). The results of the densitometry are summarized in graphs **c** and **d**. The error bars in the graphs represent standard deviations. **d**, GST pull-down assay for histone H3 mutants by GST-tagged human CIA-I(155). **e**, Spt$^-$ phenotype of yeast *asf1/cia1* mutant strains. **f**, Spt$^-$ phenotype of yeast histone mutant strains. WT, wild type; *asf1*Δ, *asf1/cia1* disruptant; and +His and −His, media with and without histidine, respectively. **g**, Positions of the yeast Asf1p/Cia1p mutations (red) causing impaired binding for the histone H3–H4 complex (left; see **c**), and Spt$^-$ phenotype (centre; see **e**). Positions of the histone H3 mutations (red) causing impaired binding for GST-tagged human CIA-I(155) and Spt$^-$ phenotype (right, see **d** and **f**). Histone H3 is shown as a blue tube.

The interaction between human CIA-I and histone H3 was analysed using a GST pull-down assay with human CIA-I and histone mutants. Among the seven human CIA-I(155) mutants—Val45(CIA)Ala, Asp54(CIA)Ala, Val94(CIA)Ala, Arg108(CIA)Ala, Tyr112(CIA)Ala, Asn114(CIA)Ala and Arg145(CIA)Ala—Asp54(CIA)Ala and Tyr112(CIA)Ala showed approximately 20–30% loss of histone H3–H4 dimer binding activity (data not shown). Asp 54(CIA) forms hydrogen bonds with Arg 129(H3) (Supplementary Table 3), and Tyr 112(CIA) interacts hydrophobically with Leu 126(H3). We then examined the reactivity of the *Saccharomyces cerevisiae* (yeast) orthologue, Asf1p/Cia1p, with histone H3–H4. The binding activity of yeast Asf1p/Cia1p(169) was more sensitive to mutation, probably reflecting the sequence difference between yeast Asf1p/Cia1p and human CIA-I. In yeast Asf1p/Cia1p(169), mutations on Val 45(Asf1), Asp 54(Asf1), Val 94(Asf1) and Tyr 112(Asf1) impaired the interaction with the histone H3–H4 dimer (Fig. 2c). We also analysed mutations of the histone H3 residues that interact with the primary binding site. Histone H3 mutants Leu126(H3)Ala and Ile130(H3)Ala showed decreased binding activities for human CIA-I(155) (Fig. 2d). Leu 126(H3) hydrophobically interacts with Val 45(CIA), Val 94(CIA) and Tyr 112(CIA) (Fig. 2a).

Next, the effects of these mutations were examined *in vivo* using the Spt phenotype in *S. cerevisiae* (Fig. 2e, f), because the *asf1/cia1* disruptant shows the Spt$^-$ phenotype[4,8]. This assay system takes advantage of transcription inhibition that is caused by insertion of a Ty or δ element at the 5′ end of a gene, disturbing its normal transcription initiation[23] (Supplementary Fig. 6). Genes encoding transcription/chromatin-related factors such as general transcription initiation factor TBP, subunits of the histone acetyltransferase complex SAGA, and some histone chaperones (Hir2 and Spt6) have been isolated using this assay[24].

We mutated most of the exposed residues on yeast Asf1p/Cia1p to identify the biologically functional surface of Asf1p/Cia1p (Supplementary Fig. 7a). The result of the assay was striking. Among the 90 mutants, 5 residues in the primary binding sites showed an Spt$^-$ phenotype (Fig. 2e). The Asp54(Asf1)Ala mutant conferred a 30–100-fold increase of the Spt$^-$ phenotype. The histone mutants Leu126(H3)Ala and Leu130(H3)Ala also showed the Spt$^-$ phenotype (Fig. 2f; Supplementary Fig. 7b). These results are remarkably consistent with those of *in vitro* analyses (Fig. 2g), suggesting that the interaction between CIA/ASF1 and histone H3 found in the complex structure is critical for cellular processes related to transcription. In addition, we recently revealed that this interaction seems to be important for both DNA replication and DNA repair[25] (M.E. and M.H., unpublished results).

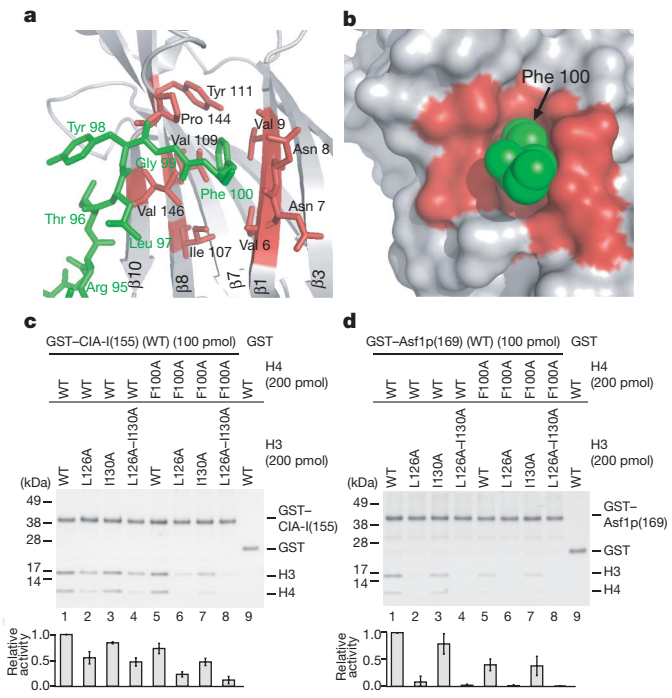


Figure 3 | Physical and functional interaction between CIA and histone H4. **a**, Interaction between the residues in the secondary binding site (red) and histone H4 (green). **b**, The hydrophobic pocket in the secondary binding site (red) of CIA-I accommodates Phe 100(H4) (green). **c**, GST pull-down assay for histone H3 and H4 mutants by GST-tagged human CIA-I(155). The results of the densitometry are summarized in graphs **c** and **d**. The error bars in the graphs represent standard deviations. **d**, GST pull-down assay for histone H3 and H4 mutants by GST-tagged yeast Asf1p/Cia1p(169).

Residues in the secondary binding site of CIA-I form an antiparallel β-sheet with histone H4 (residues 95–100) in the CIA-I–histone-H3–H4 complex (Fig. 1a). The side-chain of Phe 100(H4) is plugged into a hydrophobic cavity located between the β1- and β10-strands of CIA-I (Fig. 3a, b). The mutants in the secondary binding site, namely Phe100(H4)Ala, Tyr111(Asf1)Ala, Pro144(Asf1)Ala, Val146(Asf1)Ala, and Arg148(Asf1)Ala, showed no apparent Spt$^-$ phenotype (data not shown), suggesting a supportive role of Phe 100(H4) for the complex formation.

We then performed an *in vitro* binding assay of human CIA-I. Among the five mutants of the secondary binding site—Val6(CIA)Ala, Val9(CIA)Ala, Val109(CIA)Ala, Pro144(CIA)Ala and Val146(CIA)Ala—the Val109(CIA)Ala mutant showed an approximately 30% loss of histone binding activity (data not shown). On the histone side, the Phe100(H4)Ala mutant showed an approximately 30% loss of human CIA-I(155) binding activity. The double mutation of residues in the primary binding site and Phe 100(H4) showed a significant effect on the complex formation with CIA/ASF1 (Fig. 3c, d), suggesting that Phe 100(H4) functions in a synergetic way with residues in the primary binding site.

Our structural and biochemical analyses demonstrated for the first time that CIA-I has a histone H3–H4 tetramer-disrupting activity, which is required to form the CIA-I–histone-H3–H4 complex from the histone H3–H4 tetramer. Taking into account the results of numerous biochemical[18] and biological[6,16] analyses as well as our data, the present complex is likely to be an intermediate of the CIA-mediated nucleosome assembly/disassembly reaction and to have a critical importance in various cellular processes related to transcription, DNA replication and DNA repair[25]. Our GST pull-down assays suggested that the primary and secondary binding sites are the 'active site' of CIA/ASF1, and their synergetic action would be important for the CIA's histone H3–H4 tetramer-disrupting activity. In contrast to the nucleosome disassembly activity, the nucleosome assembly activity of CIA should be considered with other factors like DNA, because it is difficult to assemble the histone H3–H4 tetramer from the CIA–histone-H3–H4 complex owing to its stability in solution. It is of note that the above reactions could occur with a variety of histone- and CIA-interacting factors. Because the CIA-I in the CIA-I–histone-H3–H4 complex is located at nearly the same position as one of the histone H3–H4 dimers of the histone H3–H4 tetramer (Supplementary Fig. 8a), the bound CIA would cause little steric hindrance with chromatin enzymes/factors interacting with histone modifications of the histone H3–H4 dimer (Fig. 1b, Supplementary Fig. 2). Thus, CIA seems to be able to form a complex with the histone H3–H4 dimer, histone-variant-specific factor, and histone-modification-recognizing enzymes/factors (Supplementary Fig. 8b). Furthermore, CIA-variant-specific factors could interact with the CIA-I–histone-H3–H4 complex (Supplementary Fig. 8b). Nucleosome assembly and disassembly with a variety of histone- and CIA-interacting factors might be the key to the histone-modification preservation throughout the cell cycle (Supplementary Fig. 9). Functional and structural analyses of the nucleosome assembly/disassembly complex composed of the CIA-I–histone-H3–H4 complex and its interacting factors would reveal the relationship between the nucleosome assembly/disassembly regulation and histone modifications.

Note added in proof: A similar structure (yeast Asf1 in complex with the *Xenopus* H3–H4 dimer) was published[30] recently during the reviewing process of this manuscript.

METHODS

See Supplementary Information for detailed methods.

All the proteins used in the present study were overexpressed in *Escherichia coli* (Supplementary Fig. 10a–c). The CIA-I–histone-H3–H4 complex was crystallized by the hanging-drop vapour diffusion method. Diffraction data were collected at PF-AR NW12 (Tsukuba). The crystal structure was determined by the molecular replacement method using the CCP4 program suite[26]. Human CIA-I (PDB ID, 1TEY) and the histone H3–H4 dimer in the nucleosome core (PDB ID, 1KX3) were used as search models of the molecular replacement calculation.

The molecular weight of a protein was determined using a static light-scattering method using miniDAWN Tristar (Wyatt Technology). The results of the GST pull-down assays were quantified by densitometry using the program ImageJ[27]. The SDS–polyacrylamide gel electrophoresis (PAGE) gels were stained by CBB. The linearities of the densitometry with histones H3 and H4, and GST-tagged CIA/ASF1 were confirmed (Supplementary Fig. 10d–f).

In the Spt phenotype assay, threefold serial dilutions of strains with the indicated genotypes were spotted onto SC agar medium with or without histidine. The experiments were performed in duplicate and were repeated multiple times.

Received 12 October 2006; accepted 17 January 2007.
Published online 11 February 2007.

1. Le, S., Davis, C., Konopka, J. B. & Sternglanz, R. Two new S-phase-specific genes from *Saccharomyces cerevisiae. Yeast* **13,** 1029–1042 (1997).
2. Tyler, J. K. *et al.* The RCAF complex mediates chromatin assembly during DNA replication and repair. *Nature* **402,** 555–560 (1999).
3. Munakata, T., Adachi, N., Yokoyama, N., Kuzuhara, T. & Horikoshi, M. A human homologue of yeast anti-silencing factor has histone chaperone activity. *Genes Cells* **5,** 221–233 (2000).
4. Chimura, T., Kuzuhara, T. & Horikoshi, M. Identification and characterization of CIA/ASF1 as an interactor of bromodomains associated with TFIID. *Proc. Natl Acad. Sci. USA* **99,** 9334–9339 (2002).
5. Umehara, T. & Horikoshi, M. Transcription initiation factor IID-interactive histone chaperone CIA-II implicated in mammalian spermatogenesis. *J. Biol. Chem.* **278,** 35660–35667 (2003).
6. Tagami, H., Ray-Gallet, D., Almouzni, G. & Nakatani, Y. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* **116,** 51–61 (2004).
7. Green, E. M. *et al.* Replication-independent histone deposition by the HIR complex and Asf1. *Curr. Biol.* **15,** 2044–2049 (2005).
8. Sharp, J. A., Fouts, E. T., Krawitz, D. C. & Kaufman, P. D. Yeast histone deposition protein Asf1p requires Hir proteins and PCNA for heterochromatic silencing. *Curr. Biol.* **11,** 463–473 (2001).
9. Adkins, M. W., Howar, S. R. & Tyler, J. K. Chromatin disassembly mediated by the histone chaperone Asf1 is essential for transcriptional activation of the yeast *PHO5* and *PHO8* genes. *Mol. Cell* **14,** 657–666 (2004).
10. Schwabish, M. A. & Struhl, K. Asf1 mediates histone eviction and deposition during elongation by RNA polymerase II. *Mol. Cell* **22,** 415–422 (2006).
11. Emili, A., Schieltz, D. M., Yates, J. R. III & Hartwell, L. H. Dynamic interaction of DNA damage checkpoint protein Rad53 with chromatin assembly factor Asf1. *Mol. Cell* **7,** 13–20 (2001).
12. Hu, F., Alcasabas, A. & Elledge, A. J. Asf1 links Rad53 to control of chromatin assembly. *Genes Dev.* **15,** 1061–1066 (2001).
13. Osada, S. *et al.* The yeast SAS (something about silencing) protein complex contains a MYST-type putative acetyltransferase and functions with chromatin assembly factor ASF1. *Genes Dev.* **15,** 3155–3168 (2001).
14. Meijsing, S. H. & Ehrenhofer-Murray, A. E. The silencing complex SAS-I links histone acetylation to the assembly of repressed chromatin by CAF-I and Asf1 in *Saccharomyces cerevisiae. Genes Dev.* **15,** 3169–3182 (2001).
15. Umehara, T., Chimura, T., Ichikawa, N. & Horikoshi, M. Polyanionic stretch-deleted histone chaperone cia1/Asf1p is functional both *in vivo* and *in vitro. Genes Cells* **7,** 59–73 (2002).
16. Groth, A. *et al.* Human Asf1 regulates the flow of S phase histones during replicational stress. *Mol. Cell* **17,** 301–311 (2005).
17. Sillje, H. H. & Nigg, E. A. Identification of human Asf1 chromatin assembly factors as substrates of Tousled-like kinases. *Curr. Biol.* **11,** 1068–1073 (2001).
18. English, C. M., Maluf, N. K., Tripet, B., Churchill, M. E. & Tyler, J. K. ASF1 binds to a heterodimer of histones H3 and H4: a two-step mechanism for the assembly of the H3–H4 heterotetramer on DNA. *Biochemistry* **44,** 13673–13682 (2005).
19. Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* **403,** 41–45 (2000).
20. Cosgrove, M. S., Boeke, J. D. & Wolberger, C. Regulated nucleosome mobility and the histone code. *Nature Struct. Mol. Biol.* **11,** 1037–1043 (2004).
21. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389,** 251–260 (1997).
22. Mousson, F. *et al.* Structural basis for the interaction of Asf1 with histone H3 and its functional implications. *Proc. Natl Acad. Sci. USA* **102,** 5975–5980 (2005).
23. Silverman, S. J. & Fink, G. R. Effects of Ty insertions on *HIS4* transcription in *Saccharomyces cerevisiae. Mol. Cell. Biol.* **4,** 1246–1251 (1984).
24. Yamaguchi, Y., Narita, T., Inukai, N., Wada, T. & Handa, H. *SPT* genes: key players in the regulation of transcription, chromatin structure and other cellular processes. *J. Biochem.* **129,** 185–191 (2001).
25. Matsubara, K., Sano, N., Umehara, T. & Horikoshi, M. Global analysis of functional surfaces of core histones with comprehensive point mutants. *Genes Cells* **12,** 13–33 (2007).
26. Collaborative computational project number 4. CCP4 suite: Programs for protein crystallography by collaborative computational project number 4. *Acta Crystallogr. D.* **50,** 760–763 (1994).
27. Abramoff, M. D., Magelhaes, P. J. & Ram, S. J. Image processing with ImageJ. *Biophotonics International* **11,** 36–42 (2004).
28. Tang, Y. *et al.* Structure of a human ASF1–HIRA complex and insights into specificity of histone chaperone complex assembly. *Nature Struct. Mol. Biol.* **13,** 921–929 (2006).
29. DeLano, W. L. The PyMOL Molecular Graphics System ⟨http://www.pymol.org⟩ (2006).
30. English, C. M., Adkins, M. W., Carson, J. J., Churchill, M. E. & Tyler, J. K. Structure basis of the histone chaperone activity of Asf1. *Cell* **127,** 495–508 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank T. K. Kundu for the expression vector of *Xenopus* histones; F. Winston for yeast strains; K. Matsubara, Y. Ikejiri, S. Okano and S. Yoshihara for the construction of the histone mutants; and K. Hasegawa for critical reading of the manuscript. This study was supported in part by the New Energy and Industrial Technology Development Organization (NEDO) of Japan, the Exploratory Research for Advanced Technology (ERATO) of the Japan Science and Technology Agency (JST), and Ministry of Education, Culture, Sports, Science and Technology of Japan.

Author Contributions R.N. and M.E. contributed equally to this work. R.N. and Y.A. performed crystallographic and biochemical works. M.E. and N.S. performed biochemical works and yeast genetics. T.S. performed crystallographic work. M.H. and T.S. contributed to the idea, strategy, project management and writing of the manuscript. All authors discussed the results and commented on the manuscript.

Author Information The atomic coordinates have been deposited in the Protein Data Bank (PDB code, 2IO5). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to T.S. (tsenda@jbirc.aist.go.jp) or M.H. (horikosh@iam.u-tokyo.ac.jp).

CORRIGENDUM

doi:10.1038/nature05274

Happy centenary, photon

Anton Zeilinger, Gregor Weihs, Thomas Jennewein & Markus Aspelmeyer

Nature 433, 230–238 (2005)

In the legend to Figure 1, the experiment shown was wrongly attributed to Clauser. The legend should have read 'Principle of Grangier, Roger and Aspect's experiment…(ref. 10)'. In contrast, the Clauser experiment (ref. 4) involved one beam splitter on each side with detectors in each of the resulting four output ports. Four characteristic correlations were measured. In both the Clauser (ref. 4) and the Grangier, Roger and Aspect (ref. 10) experiments the observed correlations cannot be explained via classical light fields, but can easily be understood by assuming single photons that can only be detected once behind a beam splitter.

CORRIGENDUM

doi:10.1038/nature05641

The receptors and coding logic for bitter taste

K. L. Mueller, M. A. Hoon, I. Erlenbach, J. Chandrashekar, C. S. Zuker & N. J. P. Ryba

Nature 434, 225–229 (2005)

C.S.Z., N.J.P.R., K.L.M. and M.A.H. filed a patent application relevant to this work on 10 September 1999 (patent number US6558910), which should therefore have been declared as a competing financial interest.

CORRIGENDUM

doi:10.1038/nature05686

Half-metallic graphene nanoribbons

Young-Woo Son, Marvin L. Cohen & Steven G. Louie

Nature 444, 347–349 (2006)

In Fig. 2b of this Letter, the contour values were incorrectly normalized. The maximum and minimum values of ± 1.4 in the scale bar in Fig. 2b should read ± 36.6. This error does not affect any of our results. We thank E. Rudberg for pointing out this error.

CORRIGENDUM

doi:10.1038/nature05606

The prolyl isomerase Pin1 regulates amyloid precursor protein processing and amyloid-β production

L. Pastorino, A. Sun, P.-J. Lu, X. Z. Zhou, M. Balastik, G. Finn, G. Wulf, J. Lim, S.-H. Li, X. Li, W. Xia, L. K. Nicholson & K. P. Lu

Nature 440, 528–534 (2006)

During editing to meet *Nature*'s limits on length, we removed a reference to an earlier paper[1] reporting that the prolyl isomerase Pin1 promotes production of Alzheimer's amyloid-β (Aβ) from β-cleaved amyloid precursor protein (APP). That paper reported that Pin1 did not bind to full-length APP, but rather to the phosphorylated Thr 668–Pro motif of the carboxy-terminal C99 fragment of APP; Aβ production in Pin1-knockout mice was reduced only from this fragment.

1. Akiyama, H., Shin, R. W., Uchida, C., Kitamoto, T. & Uchida, T. Prolyl isomerase Pin1 facilitates production of Alzheimer's amyloid-β from β-cleaved amyloid precursor protein *Biochem. Biophys. Res. Commun.* 336, 521–529 (2005).

CORRIGENDUM

doi:10.1038/nature05608

Human embryonic stem cell lines derived from single blastomeres

Irina Klimanskaya, Young Chung, Sandy Becker, Shi-Jiang Lu & Robert Lanza

Nature 444, 481–485 (2006); doi:10.1038/nature05142 and *Nature* 444, 512 (2006); doi:10.1038/nature05366

The last sentence of the penultimate paragraph of this Letter should read "Notably, individual morula (8–16 cell)-stage blastomeres have not been shown to have the intrinsic capacity to generate a complete organism in most mammalian species." (see refs 1 and 2).

1. Moore, N. W., Adams, C. E. & Rowson, L. E. A. Developmental potential of single blastomeres of the rabbit egg. *J. Reprod. Fertil.* 17, 527–531 (1968).
2. Willadsen, S. M. The developmental capacity of blastomeres from four and eight-cell sheep embryos. *J. Embryol. Exp. Morph.* 65, 165–172 (1981).

CORRIGENDUM

doi:10.1038/nature05274

Happy centenary, photon

Anton Zeilinger, Gregor Weihs, Thomas Jennewein & Markus Aspelmeyer

Nature 433, 230–238 (2005)

In the legend to Figure 1, the experiment shown was wrongly attributed to Clauser. The legend should have read 'Principle of Grangier, Roger and Aspect's experiment…(ref. 10)'. In contrast, the Clauser experiment (ref. 4) involved one beam splitter on each side with detectors in each of the resulting four output ports. Four characteristic correlations were measured. In both the Clauser (ref. 4) and the Grangier, Roger and Aspect (ref. 10) experiments the observed correlations cannot be explained via classical light fields, but can easily be understood by assuming single photons that can only be detected once behind a beam splitter.

CORRIGENDUM

doi:10.1038/nature05641

The receptors and coding logic for bitter taste

K. L. Mueller, M. A. Hoon, I. Erlenbach, J. Chandrashekar, C. S. Zuker & N. J. P. Ryba

Nature 434, 225–229 (2005)

C.S.Z., N.J.P.R., K.L.M. and M.A.H. filed a patent application relevant to this work on 10 September 1999 (patent number US6558910), which should therefore have been declared as a competing financial interest.

CORRIGENDUM

doi:10.1038/nature05686

Half-metallic graphene nanoribbons

Young-Woo Son, Marvin L. Cohen & Steven G. Louie

Nature 444, 347–349 (2006)

In Fig. 2b of this Letter, the contour values were incorrectly normalized. The maximum and minimum values of ± 1.4 in the scale bar in Fig. 2b should read ± 36.6. This error does not affect any of our results. We thank E. Rudberg for pointing out this error.

CORRIGENDUM

doi:10.1038/nature05606

The prolyl isomerase Pin1 regulates amyloid precursor protein processing and amyloid-β production

L. Pastorino, A. Sun, P.-J. Lu, X. Z. Zhou, M. Balastik, G. Finn, G. Wulf, J. Lim, S.-H. Li, X. Li, W. Xia, L. K. Nicholson & K. P. Lu

Nature 440, 528–534 (2006)

During editing to meet *Nature*'s limits on length, we removed a reference to an earlier paper[1] reporting that the prolyl isomerase Pin1 promotes production of Alzheimer's amyloid-β (Aβ) from β-cleaved amyloid precursor protein (APP). That paper reported that Pin1 did not bind to full-length APP, but rather to the phosphorylated Thr 668–Pro motif of the carboxy-terminal C99 fragment of APP; Aβ production in Pin1-knockout mice was reduced only from this fragment.

1. Akiyama, H., Shin, R. W., Uchida, C., Kitamoto, T. & Uchida, T. Prolyl isomerase Pin1 facilitates production of Alzheimer's amyloid-β from β-cleaved amyloid precursor protein *Biochem. Biophys. Res. Commun.* 336, 521–529 (2005).

CORRIGENDUM

doi:10.1038/nature05608

Human embryonic stem cell lines derived from single blastomeres

Irina Klimanskaya, Young Chung, Sandy Becker, Shi-Jiang Lu & Robert Lanza

Nature 444, 481–485 (2006); doi:10.1038/nature05142 and *Nature* 444, 512 (2006); doi:10.1038/nature05366

The last sentence of the penultimate paragraph of this Letter should read "Notably, individual morula (8–16 cell)-stage blastomeres have not been shown to have the intrinsic capacity to generate a complete organism in most mammalian species." (see refs 1 and 2).

1. Moore, N. W., Adams, C. E. & Rowson, L. E. A. Developmental potential of single blastomeres of the rabbit egg. *J. Reprod. Fertil.* 17, 527–531 (1968).
2. Willadsen, S. M. The developmental capacity of blastomeres from four and eight-cell sheep embryos. *J. Embryol. Exp. Morph.* 65, 165–172 (1981).

CORRIGENDUM

doi:10.1038/nature05274

Happy centenary, photon

Anton Zeilinger, Gregor Weihs, Thomas Jennewein & Markus Aspelmeyer

Nature 433, 230–238 (2005)

In the legend to Figure 1, the experiment shown was wrongly attributed to Clauser. The legend should have read 'Principle of Grangier, Roger and Aspect's experiment…(ref. 10)'. In contrast, the Clauser experiment (ref. 4) involved one beam splitter on each side with detectors in each of the resulting four output ports. Four characteristic correlations were measured. In both the Clauser (ref. 4) and the Grangier, Roger and Aspect (ref. 10) experiments the observed correlations cannot be explained via classical light fields, but can easily be understood by assuming single photons that can only be detected once behind a beam splitter.

CORRIGENDUM

doi:10.1038/nature05641

The receptors and coding logic for bitter taste

K. L. Mueller, M. A. Hoon, I. Erlenbach, J. Chandrashekar, C. S. Zuker & N. J. P. Ryba

Nature 434, 225–229 (2005)

C.S.Z., N.J.P.R., K.L.M. and M.A.H. filed a patent application relevant to this work on 10 September 1999 (patent number US6558910), which should therefore have been declared as a competing financial interest.

CORRIGENDUM

doi:10.1038/nature05686

Half-metallic graphene nanoribbons

Young-Woo Son, Marvin L. Cohen & Steven G. Louie

Nature 444, 347–349 (2006)

In Fig. 2b of this Letter, the contour values were incorrectly normalized. The maximum and minimum values of ± 1.4 in the scale bar in Fig. 2b should read ± 36.6. This error does not affect any of our results. We thank E. Rudberg for pointing out this error.

CORRIGENDUM

doi:10.1038/nature05606

The prolyl isomerase Pin1 regulates amyloid precursor protein processing and amyloid-β production

L. Pastorino, A. Sun, P.-J. Lu, X. Z. Zhou, M. Balastik, G. Finn, G. Wulf, J. Lim, S.-H. Li, X. Li, W. Xia, L. K. Nicholson & K. P. Lu

Nature 440, 528–534 (2006)

During editing to meet *Nature*'s limits on length, we removed a reference to an earlier paper[1] reporting that the prolyl isomerase Pin1 promotes production of Alzheimer's amyloid-β (Aβ) from β-cleaved amyloid precursor protein (APP). That paper reported that Pin1 did not bind to full-length APP, but rather to the phosphorylated Thr 668–Pro motif of the carboxy-terminal C99 fragment of APP; Aβ production in Pin1-knockout mice was reduced only from this fragment.

1. Akiyama, H., Shin, R. W., Uchida, C., Kitamoto, T. & Uchida, T. Prolyl isomerase Pin1 facilitates production of Alzheimer's amyloid-β from β-cleaved amyloid precursor protein *Biochem. Biophys. Res. Commun.* 336, 521–529 (2005).

CORRIGENDUM

doi:10.1038/nature05608

Human embryonic stem cell lines derived from single blastomeres

Irina Klimanskaya, Young Chung, Sandy Becker, Shi-Jiang Lu & Robert Lanza

Nature 444, 481–485 (2006); doi:10.1038/nature05142 and *Nature* 444, 512 (2006); doi:10.1038/nature05366

The last sentence of the penultimate paragraph of this Letter should read "Notably, individual morula (8–16 cell)-stage blastomeres have not been shown to have the intrinsic capacity to generate a complete organism in most mammalian species." (see refs 1 and 2).

1. Moore, N. W., Adams, C. E. & Rowson, L. E. A. Developmental potential of single blastomeres of the rabbit egg. *J. Reprod. Fertil.* 17, 527–531 (1968).
2. Willadsen, S. M. The developmental capacity of blastomeres from four and eight-cell sheep embryos. *J. Embryol. Exp. Morph.* 65, 165–172 (1981).

CORRIGENDUM

doi:10.1038/nature05274

Happy centenary, photon

Anton Zeilinger, Gregor Weihs, Thomas Jennewein & Markus Aspelmeyer

Nature 433, 230–238 (2005)

In the legend to Figure 1, the experiment shown was wrongly attributed to Clauser. The legend should have read 'Principle of Grangier, Roger and Aspect's experiment…(ref. 10)'. In contrast, the Clauser experiment (ref. 4) involved one beam splitter on each side with detectors in each of the resulting four output ports. Four characteristic correlations were measured. In both the Clauser (ref. 4) and the Grangier, Roger and Aspect (ref. 10) experiments the observed correlations cannot be explained via classical light fields, but can easily be understood by assuming single photons that can only be detected once behind a beam splitter.

CORRIGENDUM

doi:10.1038/nature05641

The receptors and coding logic for bitter taste

K. L. Mueller, M. A. Hoon, I. Erlenbach, J. Chandrashekar, C. S. Zuker & N. J. P. Ryba

Nature 434, 225–229 (2005)

C.S.Z., N.J.P.R., K.L.M. and M.A.H. filed a patent application relevant to this work on 10 September 1999 (patent number US6558910), which should therefore have been declared as a competing financial interest.

CORRIGENDUM

doi:10.1038/nature05686

Half-metallic graphene nanoribbons

Young-Woo Son, Marvin L. Cohen & Steven G. Louie

Nature 444, 347–349 (2006)

In Fig. 2b of this Letter, the contour values were incorrectly normalized. The maximum and minimum values of ± 1.4 in the scale bar in Fig. 2b should read ± 36.6. This error does not affect any of our results. We thank E. Rudberg for pointing out this error.

CORRIGENDUM

doi:10.1038/nature05606

The prolyl isomerase Pin1 regulates amyloid precursor protein processing and amyloid-β production

L. Pastorino, A. Sun, P.-J. Lu, X. Z. Zhou, M. Balastik, G. Finn, G. Wulf, J. Lim, S.-H. Li, X. Li, W. Xia, L. K. Nicholson & K. P. Lu

Nature 440, 528–534 (2006)

During editing to meet *Nature*'s limits on length, we removed a reference to an earlier paper[1] reporting that the prolyl isomerase Pin1 promotes production of Alzheimer's amyloid-β (Aβ) from β-cleaved amyloid precursor protein (APP). That paper reported that Pin1 did not bind to full-length APP, but rather to the phosphorylated Thr 668–Pro motif of the carboxy-terminal C99 fragment of APP; Aβ production in Pin1-knockout mice was reduced only from this fragment.

1. Akiyama, H., Shin, R. W., Uchida, C., Kitamoto, T. & Uchida, T. Prolyl isomerase Pin1 facilitates production of Alzheimer's amyloid-β from β-cleaved amyloid precursor protein *Biochem. Biophys. Res. Commun.* 336, 521–529 (2005).

CORRIGENDUM

doi:10.1038/nature05608

Human embryonic stem cell lines derived from single blastomeres

Irina Klimanskaya, Young Chung, Sandy Becker, Shi-Jiang Lu & Robert Lanza

Nature 444, 481–485 (2006); doi:10.1038/nature05142 and *Nature* 444, 512 (2006); doi:10.1038/nature05366

The last sentence of the penultimate paragraph of this Letter should read "Notably, individual morula (8–16 cell)-stage blastomeres have not been shown to have the intrinsic capacity to generate a complete organism in most mammalian species." (see refs 1 and 2).

1. Moore, N. W., Adams, C. E. & Rowson, L. E. A. Developmental potential of single blastomeres of the rabbit egg. *J. Reprod. Fertil.* 17, 527–531 (1968).
2. Willadsen, S. M. The developmental capacity of blastomeres from four and eight-cell sheep embryos. *J. Embryol. Exp. Morph.* 65, 165–172 (1981).

CORRIGENDUM

doi:10.1038/nature05274

Happy centenary, photon

Anton Zeilinger, Gregor Weihs, Thomas Jennewein & Markus Aspelmeyer

Nature 433, 230–238 (2005)

In the legend to Figure 1, the experiment shown was wrongly attributed to Clauser. The legend should have read 'Principle of Grangier, Roger and Aspect's experiment…(ref. 10)'. In contrast, the Clauser experiment (ref. 4) involved one beam splitter on each side with detectors in each of the resulting four output ports. Four characteristic correlations were measured. In both the Clauser (ref. 4) and the Grangier, Roger and Aspect (ref. 10) experiments the observed correlations cannot be explained via classical light fields, but can easily be understood by assuming single photons that can only be detected once behind a beam splitter.

CORRIGENDUM

doi:10.1038/nature05641

The receptors and coding logic for bitter taste

K. L. Mueller, M. A. Hoon, I. Erlenbach, J. Chandrashekar, C. S. Zuker & N. J. P. Ryba

Nature 434, 225–229 (2005)

C.S.Z., N.J.P.R., K.L.M. and M.A.H. filed a patent application relevant to this work on 10 September 1999 (patent number US6558910), which should therefore have been declared as a competing financial interest.

CORRIGENDUM

doi:10.1038/nature05686

Half-metallic graphene nanoribbons

Young-Woo Son, Marvin L. Cohen & Steven G. Louie

Nature 444, 347–349 (2006)

In Fig. 2b of this Letter, the contour values were incorrectly normalized. The maximum and minimum values of ± 1.4 in the scale bar in Fig. 2b should read ± 36.6. This error does not affect any of our results. We thank E. Rudberg for pointing out this error.

CORRIGENDUM

doi:10.1038/nature05606

The prolyl isomerase Pin1 regulates amyloid precursor protein processing and amyloid-β production

L. Pastorino, A. Sun, P.-J. Lu, X. Z. Zhou, M. Balastik, G. Finn, G. Wulf, J. Lim, S.-H. Li, X. Li, W. Xia, L. K. Nicholson & K. P. Lu

Nature 440, 528–534 (2006)

During editing to meet *Nature*'s limits on length, we removed a reference to an earlier paper[1] reporting that the prolyl isomerase Pin1 promotes production of Alzheimer's amyloid-β (Aβ) from β-cleaved amyloid precursor protein (APP). That paper reported that Pin1 did not bind to full-length APP, but rather to the phosphorylated Thr 668–Pro motif of the carboxy-terminal C99 fragment of APP; Aβ production in Pin1-knockout mice was reduced only from this fragment.

1. Akiyama, H., Shin, R. W., Uchida, C., Kitamoto, T. & Uchida, T. Prolyl isomerase Pin1 facilitates production of Alzheimer's amyloid-β from β-cleaved amyloid precursor protein *Biochem. Biophys. Res. Commun.* 336, 521–529 (2005).

CORRIGENDUM

doi:10.1038/nature05608

Human embryonic stem cell lines derived from single blastomeres

Irina Klimanskaya, Young Chung, Sandy Becker, Shi-Jiang Lu & Robert Lanza

Nature 444, 481–485 (2006); doi:10.1038/nature05142 and *Nature* 444, 512 (2006); doi:10.1038/nature05366

The last sentence of the penultimate paragraph of this Letter should read "Notably, individual morula (8–16 cell)-stage blastomeres have not been shown to have the intrinsic capacity to generate a complete organism in most mammalian species." (see refs 1 and 2).

1. Moore, N. W., Adams, C. E. & Rowson, L. E. A. Developmental potential of single blastomeres of the rabbit egg. *J. Reprod. Fertil.* 17, 527–531 (1968).
2. Willadsen, S. M. The developmental capacity of blastomeres from four and eight-cell sheep embryos. *J. Embryol. Exp. Morph.* 65, 165–172 (1981).

naturejobs

THE CAREERS MAGAZINE FOR SCIENTISTS

Are blogs the new way to get a job? With 55 million blogs on the Internet and 100,000 more being created daily, some bloggers are wondering whether personal weblogs will replace traditional CVs and résumés when job-hunting. Rod Boothby, the author of Innovation Creators (www.innovationcreators.com), blogged about how his online entries had helped to advance his career. He wrote that blogs can give prospective employers extra information about you — in particular, they can show how you interact with people, at least online, and how you handle criticism, praise and complaints.

That may give some of the thousands of postdoc and graduate-student bloggers reason for pause. Many hide behind pseudonyms to air their complaints. For instance, 'iknownotwhattodo' (formerly scienceslave) welcomes readers to "the real world of bioscience, cut throat, dirty, sleazy, full of losers, power mongers, megalomaniacs, idiots, chancers" (http://geneticredundancy.blogspot.com). That comment may well draw nods of recognition from other frustrated postdocs and students. But it's not the kind of thing you necessarily want a potential employer to read. Remember: pseudonyms can be traced — and using one might counteract any positive contributions you make to the blogosphere.

Others seek more of a balance, such as the cell-biologist postdoc author of The Daily Transcript (http://scienceblogs.com/transcript), who mentions other blogs that detail "the woes of postdoc-hood" as well as what it takes to be a pioneering scientist. Apart from linking to both, the blog expands on the second, discussing the "fine line between doggedness and dogma".

Several science bloggers write about setting up or participating in 'carnivals', which are groups of related blogs. Participating in community-based forums seems a more effective use of the medium than the solipsism of complaint. Blogs may not replace résumés or CVs, but they could — intentionally or not — show a prospective employer a side of you that you wouldn't voluntarily reveal in an interview. So perhaps it's best to treat your blog as a public document that you wouldn't mind your next boss reading.

Paul Smaglik, *Naturejobs* editor

CONTACTS
Editor: Paul Smaglik
Assistant Editor: Gene Russo

European Head Office, London
The Macmillan Building,
4 Crinan Street,
London N1 9XW, UK
Tel: +44 (0) 20 7843 4961
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

European Sales Manager:
Andy Douglas (4975)
e-mail: a.douglas@nature.com
Business Development Manager:
Amelie Pequignot (4974)
e-mail: a.pequignot@nature.com
***Nature*events:**
Claudia Paulsen Young
(+44 (0) 20 7014 4015)
e-mail: c.paulsenyoung@nature.com

France/Switzerland/Belgium:
Muriel Lestringuez (4994)
UK/Ireland/Italy/RoW:
Nils Moeller (4953)
Scandinavia/Spain/Portugal:
Evelina Rubio-Morgan (4973)
Germany/Austria/The Netherlands:
Reya Silao (4970)
Online Job Postings:
Matthew Ward (+44 (0) 20 7014 4059)

Advertising Production Manager:
Stephen Russell
To send materials use London address above.
Tel: +44 (0) 20 7843 4816
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com
***Naturejobs* web development:**
Tom Hancock
***Naturejobs* online production:**
Catherine Alexander

US Head Office, New York
75 Varick Street,
9th Floor,
New York,
NY 10013-1917
Tel: +1 800 989 7718
Fax: +1 800 989 7103
e-mail: naturejobs@natureny.com

US Sales Manager: Peter Bless

Japan Head Office, Tokyo
Chiyoda Building,
2-37 Ichigayatamachi,
Shinjuku-ku,
Tokyo 162-0843
Tel: +81 3 3267 8751
Fax: +81 3 3267 8746

Asia-Pacific Sales Manager:
Ayako Watanabe
e-mail: a.watanabe@natureasia.com

NATURE|Vol 446|15 March 2007

MOVERS

Meg Urry, chair, physics department, Yale University, New Haven, Connecticut

2001–07: Professor of physics and astronomy, and director of Yale Center for Astronomy and Astrophysics, Yale University, New Haven, Connecticut
1996–2001: Astronomer and head of the science programme selection office, Space Telescope Science Institute, Baltimore, Maryland

Meg Urry says she wasn't born to be an astronomer. Interested in everything from English to politics, she found college physics frustrating initially. But that unfamiliar experience pushed her to capitalize on the aspects she liked — the logic and reductive scientific approach — and graduate from Tufts University with a double major in physics and mathematics.

A summer job at an astronomy observatory set her on an uncharted course. "There were no women in front of me. I didn't see a mapped-out path," she says. "So I moved along with a short-term focus on my work."

Still not convinced that astronomy was for her, she pursued a physics PhD at Johns Hopkins University, but ultimately conducted her research at nearby Goddard Space Flight Center in Greenbelt, Maryland. There, Urry took advantage of a newly launched powerful X-ray satellite to conduct her PhD on blazars, highly variable energy sources associated with a galaxy's central black hole. She crafted a model to explain why these previously misunderstood objects seemed so bright and so variable.

Urry went on to the Massachusetts Institute of Technology's Center for Space Research for a postdoc working with X-ray observations of supermassive black holes, and became an expert on active galaxies, which emit energies across the electromagnetic spectrum. Next, she did a postdoc at the Space Telescope Science Institute (STScI) in Baltimore, Maryland, which runs the Hubble Space Telescope for NASA. As an STScI faculty member, she used Hubble to show that all normal galaxies have a black hole at their centre and that active galaxies are simply a phase of galactic evolution. While coordinating Hubble proposal reviews, Urry discovered her administrative bent.

Eager to teach, she became the first tenured woman in the department of physics at Yale University in 2001. She was also appointed co-chair of the National Academy of Sciences Committee on Astronomy and Astrophysics, where she helps guide strategic planning. This July, Urry will become the first woman chair of physics at Yale.

As department chair, she will encourage undergraduates to take, and more importantly enjoy, physics. She hopes to continue Yale's excellence in physics by anticipating future needs and growth areas, her primary challenge according to her predecessor, Ramamurti Shankar. "The real trick," says Urry, the once-reluctant astronomer, "is to think about what questions will be asked in ten years." ■

Virginia Gewin

BRICKS & MORTAR

Colleagues against cancer

The United Kingdom's largest new cancer research initiative for 40 years opened on 2 February. The £50-million (US$98-million) Cancer Research UK Cambridge Research Institute (CRI) is set to transform UK cancer research. Its strategic location and structure are designed to bridge basic research and clinical practice.

Built on the University of Cambridge campus, the CRI sits nextdoor to Addenbrooke's teaching hospital. With about half of the 30 group-leader posts filled, a second round of recruitment is under way. The goal is to have basic cancer-biology experts working with clinician-scientists and leaders in genomics, imaging and bioinformatics. Investigators are eager to find basic scientists with skills in epigenetics and tumour microenvironments — but are open to any outstanding candidates.

Director Bruce Ponder and his colleagues are looking for people to head specific cancer programmes such as breast, lung or prostate. They also need an epidemiologist to bring together cohorts for studies of early diagnostic intervention. Above all else, however, they want investigators eager to collaborate in the clinic.

By providing research opportunities for clinicians, deputy director and stem-cell biologist Fiona Watt hopes to help basic scientists see cancer as a disease rather than an intellectual problem. "I think sometimes, perhaps more so in Britain than the United States, there's a feeling that clinicians are not doing high-quality science," says Watt. "We've disproved that myth in one stroke."

Jason Carroll, a hormone-receptor expert and one of several junior faculty members recruited last year from the United States, says he was lured by core funding and the ease of conducting research straight away. Instead of struggling for grants in the United States, he receives funding for three postdoc or technician positions as well as for one student every two years. (Senior faculty members receive six paid positions as well as a student every two years.)

Interdisciplinary research projects include a virtual Cambridge Cancer Centre, which will bring together engineers and cell biologists. Ponder says a new master's degree will bring maths students in as well. "The idea is that our new institute will be a nucleus of massive cancer-related activity across the university," says Watt.

Once immediate talent needs are met, Ponder says they will decide how to expand existing themes or start new ones, such as drug development. ■

Virginia Gewin

POSTDOC JOURNAL

Where the rocks are

By the time you read this, I will be settling into my new geology postdoc position in Johannesburg. As I write, I'm still having trouble believing that my big move is actually happening. My visa is stuck in my passport, the flights are booked, and my life is slowly being squashed into as few boxes as possible. Yet part of me still doesn't believe that any of it is really happening; probably the same part of me that is still a little intimidated by it all.

We're told that expanding one's research horizons is an essential part of developing as a scientist. If you want to enjoy use of the best labs, and work with the best groups, you often have to look abroad for your new perspective.

Yet, another consideration for us geologists is getting access to that critical outcrop, where the rocks are of the right age and composition to record clues about our planet's past tectonic and climatic history. I spent my PhD in Southampton, on the south coast of England, suffering the consequences of my field area being located in the opposite hemisphere from my lab.

For my new project, if I need to collect more samples, or check some observations, I'll only have to endure a few hours' drive rather than a long-haul flight. For that, if nothing else, it's well worth overcoming a little trepidation about making a big change. ■

Chris Rowan is starting a postdoctoral fellowship in the geology department at the University of Johannesburg, South Africa.

The inside track from academia and industry

Leaks in the pipeline

Why do women remain curiously absent from the ranks of academia?

Mary Anne Holmes

Suzanne O'Connell

Family issues can cause women to abandon academia at every rung of the career ladder. Policy-makers have addressed some ways to get more women on to the lower rungs of the ladder. But solutions at the higher steps — tenure and beyond — are proving a little more difficult.

In the United States, the past 30 years have seen a dramatic rise in the number of women gaining PhDs in the fields of science, technology, engineering and mathematics, according to the National Science Foundation (NSF). In the geosciences, the proportion of PhD degrees awarded to women has increased from none in 1966 to 46% in 2003. But, according to a database held by the American Geological Institute, there are 'leaks' in the geosciences pipeline for academics — particularly in the hiring for assistant professor positions. In the field, 42% of BS/BA degree recipients, 45% of MS recipients and 39% of PhD recipients are women. But only 26% of assistant professors, 14% of tenured associate professors and 8% of full professors are women.

The biggest barrier lies in the structure of academia. Women may hesitate to apply for tenure-track jobs because they lack role models among the upper echelons. We conducted focus groups of active, employed geoscientists, including students, and found that nearly half of the women participants seriously considered leaving the geosciences at some point in their career, as opposed to only one-third of the men. The reasons for considering leaving are strikingly different between the two genders: the top two reasons for women were family issues (caring for children or elderly relatives) and problems with advisers (mostly a failure to communicate). By far and away, the main reason males considered leaving was an uncertain job market — a distant second was a tie between difficult classes and choosing the wrong sub-discipline. We think that 'problems with advisers' is a barrier that can be minimized by training junior (and willing senior) faculty members in mentorship.

Clearly women's biological clocks play a role. Apart from medicine, in what other profession is it common for careers to begin in the early to mid-thirties? A new assistant professor, with an average age of 33, is facing the most intense work period of his or her life. For women at this age, fertility declines every year while the chances of a miscarriage or conceiving a child with Down's syndrome increase. Few graduate schools have provisions for family leave. Most graduate students answer directly to a single PhD adviser, who might not allow time off for childbearing.

More universities should provide paid family leave for graduate students and faculty members. Only one-third of PhD-granting institutions provide any sort of daycare for graduate students and most have no childbirth policy. Stanford University recently took the lead and introduced an automatic institution-wide childbirth policy for graduate students that includes six weeks' paid leave. Offering high-quality, affordable campus childcare will mitigate worries that could seriously lessen students' academic productivity.

"Although overt discrimination against hiring women has mostly disappeared, unconscious biases persist."

Departments could actively recruit women and educate hiring committees. As departments often know about openings a couple of years in advance, potential candidates in broad subject areas can be identified and courted.

There should be an automatic extension to tenure so that junior female faculty members do not have to choose between children and their job. Tenured women continue to leave because of family responsibilities. In a 40-year academic career, why not allow a temporary (one to three years) part-time option? Better assistance in spousal employment would help as well. Pennsylvania State University, for example, has temporary two- to three-year spousal appointments.

Although overt discrimination against hiring women has mostly disappeared, unconscious biases persist. As noted previously in this column (L. Bornmann *Nature* **445,** 566; 2007), gender bias can influence the awarding of grants and academic prizes. Unconscious bias in hiring and promotion has also been documented (B. J. Tesch *et al. J. Am. Med. Assoc.* **273,** 1022–1025; 1995). There needs to be a concerted effort to bring this to an end.

To help explore some of these professional and structural impediments, we are convening a consortium of geoscientist academics in New England. This NSF-funded endeavour has three components: a week-long retreat to focus on writing in the absence of departmental and domestic distractions; skills workshops on topics such as strategic persuasion and negotiation; and workshops for departmental chairs to learn about unconscious bias and ways it can be overcome. With attention to these details, we hope that the science faculty will look more like the student body in 2027. ■

Mary Anne Holmes is at the University of Nebraska in Lincoln. Suzanne O'Connell is at Wesleyan University in Middletown, Connecticut.

National Institutes of Health

Virtual Career Center

at the tip of your mouse

www.training.nih.gov/careers

-EXPLORE CAREER OPTIONS
-CONTINUE YOUR EDUCATION
-SEARCH FOR A JOB

Careers in science, medicine, and beyond at the college, postbaccalaureate, graduate, and postdoctoral levels

Office of Intramural Training and Education
(800) 445-8283

Postdoctoral, Research, and Clinical Fellowships at the National Institutes of Health

www.training.nih.gov/pdopenings

www.training.nih.gov/clinopenings

Train at the bench, the bedside, or both

Office of Intramural Training and Education
Bethesda, Maryland 20892
800.445.8283

Deputy Director
Fogarty International Center (FIC)

The NIH is seeking exceptional candidates for the position of Deputy Director, Fogarty International Center (FIC) to advance health through international scientific cooperation and to promote and support scientific research and training internationally to reduce disparities in global health. This position offers a unique opportunity for the right individual to assist the Director, FIC in providing strong and visionary leadership to an organization dedicated to international health and global collaboration. This position serves as Deputy to the Director, FIC in coordination of all activities related to the mission and functions of the Center, in the development and execution of plans and policies of the FIC and in the allocation of resources. The Deputy Director provides expert advice and counsel to the Director on the development of and opportunities for international research and research capacity building; on the development and coordination of international agreements in which the NIH participates; and on the development and dissemination of information, nationally and internationally, on the Center's international research and training initiatives. Applicants may browse the FIC Home Page at **http://www.fic.nih.gov**. Applicants must possess an M.D., Ph.D., or equivalent degree and be an experienced administrator with a broad international programmatic or scientific background; able to interact with full authority; have the demonstrated capability to plan and direct programs of international importance; and have the ability to communicate with and obtain cooperation of the public, private and international organizations and individuals from the international sector. Salary is commensurate with his/her qualifications and experience. Full Federal benefits including leave, health and life insurance, long-term care insurance, retirement, and savings plan (401k equivalent) will be provided. A detailed vacancy announcement that includes application procedures is available at: **http://www.jobs.nih.gov (under vacancy announcement FIC-07-174209)**. Questions may be addressed to **Ms. Yolette Lee, 301-594-2792 e-mail: leeyol@mail.nih.gov**. CV and bibliography, including a statement addressing the qualifications, must be received by **April 30, 2007**.

THE NIH IS DEDICATED TO BUILDING A DIVERSE COMMUNITY IN ITS TRAINING AND EMPLOYMENT PROGRAMS

OPPORTUNITIES @ NIH THE NATIONAL INSTITUTES OF HEALTH

NW98578R